

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 18:43:46 ; Search time 91 Seconds  
(without alignments)  
2776.299 Million cell updates/sec

Title: US-09-964-956-13

Perfect score: 9990

Sequence: 1 MKAMPWNWTCLLSHLLVMG.....QKLAYKLEQVITLMSLDNKK 1896

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6786	67.9	1309	23	Novel human protei
2	6764	67.7	1884	23	Mouse ischaemic co
3	3641	36.4	1945	22	Drosophila melanog
4	2710	27.1	2135	22	Human plexin prote
5	2580	25.8	1838	22	Human plexin-B2.
6	2565	25.7	1849	22	Human plexin-B1/SE
7	2521.5	25.2	1870	22	Novel human diagno
8	2471	24.7	1992	22	Human plexin-B1/SE
9	2463.5	24.7	1925	22	Human plexin-D1.
10	1838.5	18.4	1788	22	Human plexin-B3.

11	1573.5	15.8	362	22	ABB11753	Human NOV/plexin-A
12	1486	14.9	716	22	AA093421	Human polypeptide,
13	1376	13.8	484	22	AA093435	Human polypeptide,
14	1351.5	13.5	552	20	AAV13373	Amino acid sequenc
15	1351.5	13.5	552	21	AA033420	Human PRO235 prote
16	1351.5	13.5	552	21	AA024400	Human PRO235 prote
17	1351.5	13.5	552	22	AAU12337	Human PRO235 prote
18	1351.5	13.5	552	22	AA080241	Human VESPR. Homo
19	1349	13.5	1568	20	AAV13462	Viral-encoded sema
20	1349	13.5	1568	21	AA028522	Human VESPR. Homo
21	1349	13.5	1568	22	AAU00239	Viral-encoded sema
22	1349	13.5	1568	22	AAU00435	Viral encoded sema
23	1349	13.5	1568	22	AA070131	Human VESPR. Homo
24	1349	13.5	1568	22	AA031693	Amino acid sequenc
25	1275	12.8	280	23	ABG34075	Human Pro peptide
26	1076.5	10.8	199	22	AAU00018	Human Plexin A-4.
27	920	9.2	816	22	AB066230	Drosophila melanog
28	678	6.8	729	22	AAU00020	Human Plexin prote
29	653	6.5	164	22	AA093812	Human protein sequ
30	634.5	6.4	226	23	ABP41337	Human ovarian anti
31	494	4.9	179	20	AAV74169	Human prostate tum
32	484	4.8	146	21	AA039454	Gene 16 human secr
33	484	4.8	146	21	AA039455	Human secreted pro
34	471	4.7	339	21	AA057044	Human prostate can
35	427	4.3	1400	20	AAW82791	Human RON receptor
36	367	3.7	89	22	AA064598	Human brain expres
37	367	3.7	89	22	AA037545	Peptide #11582 enc
38	367	3.7	89	23	ABG46408	Human peptide enco
39	333.5	3.3	276	22	ABG04109	Novel human diagno
40	328.5	3.3	832	22	AA030818	Human gene 1 encod
41	328.5	3.3	832	23	ABG64522	Human albumin fusi
42	328.5	3.3	837	21	AAV99410	Human PRO1480 (UNQ
43	328.5	3.3	837	22	AAU29250	Human PRO polypept
44	328.5	3.3	837	22	AA066159	Protein of the inv
45	328.5	3.3	893	23	AB097964	Human protein sequ

#### ALIGNMENTS

RESULT 1  
ABB97486  
ID ABB97486 standard; Protein; 1909 AA.  
XX  
AC ABB97486;  
XX  
AC ABB97486;  
DT 27-JUN-2002 (first entry)  
XX  
DE Novel human protein SEQ ID NO: 754.  
XX  
KW Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;  
KW antiinfertility; cerebroprotective; cytosstatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
KW expressed sequenced tag.  
XX  
OS Homo sapiens.  
XX  
PN WO200222660-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 10-SEP-2001; 2001WO-US26015.  
XX  
PR 11-SEP-2000; 2000US-0659671.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX  
DR WPI; 2002-292408/33.  
XX  
DR N-PSDB; ABN32672.

PT An isolated polynucleotide for treating diseases associated with its  
 XX encoded polypeptide such as cancer and multiple sclerosis -  
 PS Claim 20; SEQ ID NO 754; 509pp; English.

XX The present invention provides the protein and coding sequences of 444  
 CC novel human proteins. These were isolated from expressed sequences tags  
 CC (ESTs). They can be used to stimulate cell growth, to regulate  
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
 CC Parkinson's disease. The present sequence is a protein of the invention.

XX Sequence 1909 AA;

Query Match 67.98; Score 6786; DB 23; Length 1909;  
 Best Local Similarity 66.98; Pred. No. 0;  
 Matches 1278; Conservative 239; Mismatches 349; Indels 44; Gaps 9;  
 23 STILLTROPAPLSOKOR-----SFVTFRGEPAE- 49  
 6 SLLLT--AFLSMEQRPWPRALVEDSRVSVLLSVWVLLAPPAGMPQFSTPHSENRDW 63  
 50 GPHLVVDERTGHIYLGAVNRIYKLSDLKVLVTHETGPDNDPKCYPPIRIVOTCNEPLT 109  
 64 TPNHLTHOGTGAVYGAINRVYKLTGNLTQVAHKTGPEEDNKSCTPPLIVQPCSEVLT 123  
 110 TTNVNMKLLIDYKENRLIACGSLYQIGCKLLRLDELFLGEPYKHKEHYLSGVNESGSV 169  
 124 LTNVNMKLLIDYKENRLIACGSLYQIGCKLLRLDELFLGEPYKHKEHYLSGVNKTGM 183  
 170 FGIVSYNLDLKLFIATAVDGKPEYPTTISRKLTKNSEADGMFAYVHDFVSMIKI 229  
 184 YGVIVRSEGEDGLFICTAVDGKQDYPTLSRKLPRDPRESSAMLDYELHDFVSSLIKI 243  
 230 PSDFTIIPDIDYVYGFSGNRYVFLTLOPEMVSPG---STTKQVYTSKLVRCKE 286  
 244 PSDTALVSHFDIPIYVYGFSGGVYFLTVOPE--TPEGVAINSAGDLFTYSRIVRLCKD 301  
 287 DTAFNSVVEVPIGERSGVEYRLQAAYLSKAGAVLGRITLGVHPDDDLFTVFSKGQKRK 346  
 302 DPKEHSVSVPPGCTRAGVEYRLQAAYLAKPGDSLQAENITSQDDVLPAIFSKGQKQY 361  
 347 MKSLDESALCIPILKQINDRIKERLQSCYRGEGTDLAWLKVDIPCSALLTTDDNFCG 406  
 362 HHPDDSALCAFPITRAINLQIKERLQSCYQEGNLELWLLGKDVQCTKAPVPIDDNFCG 421  
 407 LDMNAPLGVSMDVRGIPVFTEDRDRMTSVIAYVYKNHSLAFVGTGSKGLKIRVDGPRGN 466  
 422 LDINQPLGGTTPVEGLTLYTSRDRMTSVIAYVYKNHSLAFVGTGSKGLKIRVDGPRGN 481  
 467 ALQYETVQW--DPGFLVRLDMAFSKDHEQLYIMSERQLTRVPVSGQYQSCGECGLSGDP 525  
 482 GVQYEMSVLKGDSFILRDMAFSIDQYLYVMSEKQVTRVPVSGCEQYTTGCEGLSSGDP 541  
 526 HCGVCLVHNTCTRKERCERSKEPRFASMEKQCVRLTVPNNISQVSNVNLVLETYNVP 585  
 542 HCGWCAHLNMCSSRDCKQAWENRFAASISQCVSLAVHPSSISVSEHSLLSLVVSAP 601  
 586 ELSAGVNTCEPDLSEMDGLVGNQIQCYSPAACEVPRITENGDHVYVQLOLKSKEGTMT 645  
 602 DLSAGIACAFNLTEVEQVSGSVQICLSPGPKVP-VIPLDQDWFGLEQLRSKETGKI 660  
 646 PASTSFVYNCVSNVNSCLSCVESPVRCCHWKYKHVCTHDPKTCSPQEGRVKLPDPDQPLL 705  
 661 FVSTEFKPYNCASHQCLCLSCVNSAFRCCHWKYKYNLCTHDPKTCSPQEGRIINISDPCQLV 720  
 706 RVDKILVPEVVIKPTLTKAKNLPOQSGORGVECTILNQSGEQRVPALRFNSSSVQCCONT 765  
 721 PTEELIPVGEVKPITLTKARNLPQSGORGVECTVNIQGAIRVPALRFNSSSVQCCONS 780

QY 766 SYSEGMENINNVVELTVVNGHFNIDNPAQNVHLKCGAMRESSCGLCLKADPDPAQW 825  
 DB 781 SYQDGMIDINLAVDFAVVWNGFIIDNPDLKVLHYKCAQORSCGLCLKADKRFECGW 840  
 QY 826 CQPGQCTLRQHCPAORSQMLELSGAKSKCTNPRITRIIPVTGPREGKTIVIRENIGL 885  
 DB 841 CSGERRCTLHQHCTSPSSPWLDMSSHNKVCNPOITTEILTVSGPPEGTRVTHIGVNLGL 900  
 QY 886 EFRDIASHVKVAGVECSPLVDGYIPAQIVCEMGEAKPSQHAGFVEICVAVCRPEFMARS 945  
 DB 901 DFEIAHHVQVAGVPCPLPCEYIAEQIVCEMHALVGTTSQPVRLCIGCECKPEFMTKS 960  
 QY 946 SOLYFMTLTLSDLKPSRGMSCGTQVITITGNLHAGSNVVMFGKQPCQLFHRHRRSPSYIV 1005  
 DB 961 HQYTFNPSVLSLNPGRGSGGTWVITGHYLCAGSSVAVYLGNOTCEFGVRSMEIV 1020  
 QY 1006 C-NTTSSDEVLEMKVSVQVDRAKIHQDLVFOYEDPTIVRIEPEWSIVSGNTPIAVMGTH 1064  
 DB 1021 CVSPSSNGLGPVPSVSDRAHVDSNLQFEYIDDPVQRVRIEPEWSIASGHTPLTITGFN 1080  
 QY 1065 LDLIQNPQIRAKHGGKEHINICEVLNATEMTQAPALAGDPHQSDLTERPEEFGFILDN 1124  
 DB 1081 LDVIOEPRIRVKFKGKSVNVCKVNTTLLCLAPSLTTRDPLGDTVERPEDFGVFN 1140  
 QY 1125 VQSLILNKNFTYYPNPVFEAFGSPGILELKPPTIILKGNLIPPVAGGNVKNLYTVL 1184  
 DB 1141 VQSLIYNDTKFYYPNPTFELLSPGLDQKPGSPILKGNLCPASGG-AKLYNTVL 1199  
 QY 1185 VGEKPTVTSDVOLLCEPNLIGRHKVMARVGMGEYSPGVYIAPDSPLSIPAIVSIIV 1244  
 DB 1200 IGETPCATVSETQLLCEPNLTGQKVLVHVGVMVFGSVISDSLLTLPVIVSIAA 1259  
 QY 1245 AGGLIIFIVAVLTAAYKRESLDTLKLQMDNLSRVALECKEFAELQTDIHELT 1304  
 DB 1260 GGSULLIIVILVTAAYKRSRENDLTLKLQMDNLSRVALECKEFAELQTDIHELT 1319  
 QY 1305 SLDLQAGIPFDYRTYMRVLPFGIEDHPVLRLDLEVPYGRVEREKLKFLAQLINNKVF 1364  
 DB 1320 SLDLQSGIPYLDYRTYMRVLPFGIEDHPVLRLDLEVPYGRVEREKLKFLAQLINNKVF 1379  
 QY 1365 LLSFIRTLQSRSMRDRGNVASLIMTVLQSKLEYATDVLKQILADLIDKNLESKNHPK 1424  
 DB 1380 LLSFIRTLQSRSMRDRGNVASLIMTVLQSKLEYATDVLKQILADLIDKNLESKNHPK 1439  
 QY 1425 LLRRTESVAKMLTNWFTLLYKFLKECAGEPLFLCAIKQOMKGPIDAITGEARYS 1484  
 DB 1440 LLRRTESVAKMLTNWFTLLYKFLKECAGEPLFLCAIKQOMKGPIDAITGEARYS 1499  
 QY 1485 LSEDKLIRQQIDYKTLVLSVSPDNVANSPEVPVKILNCDTITQVKEKILDAIFKNVPCSH 1544  
 DB 1500 LSEDKLIRQQIDYKTLVLSVSPDNVANSPEVPVKILNCDTITQVKEKILDAIFKNVPCSH 1559  
 QY 1545 RPKAADMDLEWRQSGABWILQDEDTTKIENDWKRNLANTLAHVQVDPDGSVVALVSKQVTA 1604  
 DB 1560 RPKAADMDLEWRQSGABWILQDEDTTKIENDWKRNLANTLAHVQVDPDGSVVALVSKQVTA 1619  
 QY 1605 YNAVNTSVTSRTSASKYENMIRYTGSPDSLSRTPMTIPDLESQVMMHVLVKNHEDGQK 1664  
 DB 1620 YNAVNTSVTSRTSASKYENMIRYTGSPDSLSRTPMTIPDLESQVMMHVLVKNHEDGQK 1679  
 QY 1665 BGDGRSKVSEIYLRLLATKGTLOKFDVDDLPEITFSTAHRSALPLAIKYMFPDFLDEQA 1724  
 DB 1680 BGDGRSKVSEIYLRLLATKGTLOKFDVDDLPEITFSTAHRSALPLAIKYMFPDFLDEQA 1739  
 QY 1725 DKHGIDHPVHTWKSNCPLRFVWNMKNPQFVDIHKNSITDACLVSVAOTFMDSCST 1784  
 DB 1740 DRHSIHTDVRHWKSNCPLRFVWNMKNPQFVDIHKNSITDACLVSVAOTFMDSCST 1799  
 QY 1785 SEHRLGKDSPNKLLIYAKDI PSYKNWVERYSIDIGMPAISODDMNAYLAESQRMHNEF 1844  
 DB 1800 SEHRLGKDSPNKLLIYAKDI PSYKNWVERYSIDIGMPAISODDMNAYLAESQRMHNEF 1859



XX	WPI; 2001-656860/75.
DR	N-PSDB; ABL09552.
XX	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
PT	
PT	
XX	Disclosure; SEQ ID NO 23139; 21pp + Sequence Listing; English.
PS	
XX	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC	
XX	Sequence 1945 AA;
SQ	
	Query Match 36.4%; Score 3641; DB 22; Length 1945;
	Best Local Similarity 42.0%; Pred. No. 4.5e-307;
	Matches 829; Conservative 320; Mismatches 637; Indels 190; Gaps 60;
Qy	26 LTRQPAPLSOKQRSVTFPRGEPA--EGFNHVLVDERTGHIYLVGNVRNRYKLSDDLKLVT 83
Dd	48 LTANAPAIKNAKNLNSTITNVAAPDTKLNHLLDVTTIGRVFVGVNRLYQLSPDLELSET 107
Qy	84 HETGPDEDNKPYPPRIVQTCC--NEPLTNTNNVMKLLIDYKENRLIACSLYGQICKLL 141
Dd	108 VKTGPNDSVEC-----SILDCPLNAVRSPTDNYNKVLIDRATSRLLIACGSFLFGCTCVR 163
Qy	142 RLDELFKLGEPYHKHEHYLSGVNESGVFGVIYSYSLDDKLFATAVDGKPEY----- 195
Dd	164 NLQNVSII-----EHEVPD-----AVANDANSTVAFIA---PGPQHPTNVNM 205
Qy	196 -----RPTISRKLTKNSEADGMFAFYVPHDFEFVASMIKIPSDTFT---II 237
Dd	206 YVGVTYTNNSPYRSEITPAVASRSLEKTK----ME-----QIASSAVTTGRTTINSYAR 255
Qy	238 PDFDIYVVYGFSGSNFYFITLOPEMVSPPGSTTKEOYVTSKLRUCKEDTAFNSYVEVP 297
Dd	256 ETYPNVYVYGFSSERFSFYLTQ--LKSHSHSSPRE--YITKLVRIQOEDSNYYSYTEIP 311
Qy	298 IGC---ERSGVEYELLQAAYLSKAGAVLGRTLGVHPDDDLLFTVFSGOKRMKSLSDESA 354
Dd	312 VECISDRAGGTGKFMVLVOGLPKFPSSDLAQLGISIQDDVLFAVFSKGEGN--TPTNNSA 369
Qy	355 LCIFILKQINDRIKERLQSCYRGEGTLLDALWKVKDIPC--SSALLTIDDNFCGLDMNAPL 413
Dd	370 LCIIYSLASIRRFQMNIKSCFNGSMRGLDPIS--PSMPCVLTKLTQIGEDFCGLDVNSPL 428
Qy	414 GVSDMVREGIPVFTEDRDMTSVIAYVYNKHSLAFVGIKSGKLKKIRDVDGPRNALQYETV 473
Dd	429 GGETHPTTSVPY-AMFNTKLLTSVAATSTSGYTVFVFGTSDGFLFKKVVIS-SIANEYASF 486
Qy	474 QVVDPGPVLRDMAFSKDHOELYIMSEROLTRVPVESCGYQSCGCELGSGDPHCWCVLH 533
Dd	487 AVDLGSEINRDMQFDNQNLXIYVMSKTQSVKVKVFDICSDYKTCGCLCARDPYCGWCSSL 546
Qy	534 NTCRTERKC-ERSKEPRRRPAS-EMKQCVRILT-VHPNNISVSQYVNL-LVLETYINVPELSA 589
Dd	547 NKCSPRSNCQDANDPLYVWSYKTKCKTTITSVVPHQIORTTARTLELIID--HIPOLKE 604
Qy	590 GVNCTFEEDLSMDGLVWG-----NOIOCYSPAAXEVPRIIITENGDDH--VVOLQLKSKETG 643
Dd	605 NLICAF--TTEDKALFTNATKKRGNVNCITPTDMLPOI---EQGKHHFTAKLSVRTR-NG 659
Qy	644 MTPASTSFVYCNVSHNSCLSCVSPSPYRCHWKYRHVCTHD-PKTC----- 688





1916 EYLLRLLSMKGTLOKFDVDDLFQVILSTR---PVLAVKYFFDLDEQAQHGSDQPT 1972  
1735 RHTWKNCLPLRFVWNIKNPQFVDFDIHKNSITDACLSSVVAOTFMDSCSTSHRLGKOSP 1794  
1973 IHIWNTSLPLRFVWNIKNPQFVDFDIHKNSITDACLSSVVAOTFMDSCSTSHRLGKOSP 2032  
1795 SNKLLYAKDIPSKWVERYSYSDIGKMPAISQDQNNAYLAEOSRMHMFNTMSALSEIF 1854  
2033 INKLLYARDIPYKRMVERYYADIRQTPASDOEMNSVLAELSMWYSGDLGARVALHELY 2092  
1855 SYGVKYSSEILGFLDHDGCGKQKAYLEQV 1886  
2093 KYINKYDQIITALEBEDGTQAQKQVLGYRLQOI 2124

RESULT 5

AAU00015

ID AAU00015 standard; Protein; 1838 AA.

AC AAU00015;

DT 09-MAY-2001 (first entry)

XX Human Plexin-B2.

DE Human; Plexin-B2; semaphorin domain; hyperplasia; neoplasia; cancer;

KW neurodegenerative disease; autoimmune disease; lupus; multiple sclerosis;

KW inflammatory bowel disease; diabetes type I; rheumatoid arthritis;

KW immunogen; antibody; MET related sequence; MRS.

XX Homo sapiens.

OS

XX Key

PH Location/Qualifiers

FT 1..18

FT /label= "Signal peptide"

FT 1..1199

FT /label= "Extracellular domain"

FT 19..1838

FT /label= "Mature Plexin-B2"

FT 19..518

FT /label= "Semaphorin domain"

FT 451..530

FT /label= "Primary Met related sequence, MRS"

FT 601..680

FT /label= "Secondary MET related sequence, MRS"

FT 751..830

FT /label= "Tertiary MET related sequence, MRS"

FT 800..1010

FT /label= "G-P repeat region"

FT 1196..1215

FT /label= "Transmembrane domain"

XX WO200114420-A2.

PN

XX

PD 01-MAR-2001.

XX

XX 25-AUG-2000; 2000WO-US23365.

PF

XX 25-AUG-1999; 99US-0150576.

PR (UYTO-) UNIV TORINO.

XX (REGC) UNIV CALIFORNIA.

XX

PA Artigiani S, Comoglio PM, Goodman CS, Tesier-Lavigne M;

PI Tamagnone L;

XX WPI; 2001-226610/23.

DR N-PSDB; AAS00020.

XX

XX New plexin polynucleotides and polypeptides, useful in diagnosis,

PT therapy and in producing compounds for treating diseases involving

PT aberrant cell growth (e.g. cancer) or immune regulation (e.g.

PT

PT autoimmune diseases)

XX Claim 4; Page 52-56; 79pp; English.

PS

XX The sequence represents Human Plexin-B2. Plexins are large transmembrane

CC proteins whose extracellular domain shares homology with Scatter factor

CC receptors and contain an approximately 500 amino acid semaphorin domain.

CC The plexin polynucleotides and polypeptides, and plexin-specific binding

CC agents are useful in diagnosis, therapy and in the biopharmaceutical

CC industry. In particular, the plexin polynucleotides and polypeptides are

CC useful for generating compounds (e.g. plexin-specific binding agents or

CC antibodies) for treating or diagnosing a disease or disorder involving

CC aberrant cell growth (e.g. hyperplasia, neoplasia, cancer or

CC neurodegenerative disease), or diseases or disorders involving aberrant

CC immune regulation (e.g. autoimmune diseases such as lupus, inflammatory

CC bowel disease or Diabetes Type I), or immunosuppressive diseases such as

CC multiple sclerosis or rheumatoid arthritis.

XX

SQ Sequence 1838 AA;

Query Match 25.8%; Score 2580; DB 22; Length 1838;

Best Local Similarity 33.2%; Pred. No. 1.5e-214;

Matches 65%; Conservative 344; Mismatches 713; Indels 260; Gaps 61;

24 TLLTROPAPLSQKQSFVTRGEPAGFNHLLVVDERTGHIYLGAVNNRIYKSLDKVLVT 83

9 TLLGLLGGAGSLRPRKLDFFRSE--KELNHLAVDEASGVVYLGAVNALYQDKAKLQEQ 66

84 HETGPDENPKCVPRIVQTCNEPLTTNNNNKMLLDYKENRLIAGSLYQGIKLLRL 143

67 VATGPALDNKKCTPPTTEASQCHE-ABMTDNVNQLLLDPPRKLVEGSGFKGICALRAL 125

144 ED--LFLGEPYHKEHYLGVNNGSVFGVIVSYNLDLDFATAVDGKPEYFPFTISS 201

126 SNISLFLFYEDSGSEKSFVANSDEGATVGLVSGTGGDRVLFGKNGPHDNGIIVST 185

202 RKLTKNSEADGMPAYVYHDFVNASMIKIPSDYFTIIPDFDIYVYVYVYVYVYVYVYVY 261

186 RLIDRTDSREAFAYTDHATYKAGYLSLTNTQF-----VAAPEDGPFYVFFVFNQ 235

262 EMVSPGSTTKQVYTSKLVRLCKEDTAFNYSYVEVPIGCSRGVEYELLQAAVLSKAGAV 321

236 DK-HPARNRT-----LLARMCREDPNYSYLEMDLQCRDPIH-----AAAF 279

322 LGRTLVGHPDDLLFTVFSKQGRKMSLDESALCTIFILKQINDRIKERLQSCYRGEGTL 381

280 LAASVAAPSGRVLVAVFRDS--RSSGGPAGLCFLPDKVHAKMEANRACY--TGTR 335

382 DLAWLKVK---DIPCSSALLTTIDDNF-CGLD-MNAPLGVSMDVMVRGIPVFTEDRDMTSV 435

336 EARDIFYKPFHGDICQGGHAPGSSKSPFGSGEHLPPYLGSRDGLRGTAVLQRGGLNLTA 395

436 IAYVYKXNSLAFVGTSGKLLKRVDPGRGNALQYETVQVDPGVLRDMAFSKDHQOLY 495

396 TVAAENNHHTVAFVLTGSDGRILKYVLT-PDGTSTSEYDSILVEINKRVKRDVLVSGDLG 454

496 IMSEKLTVPVSVESCGYOSCGECLSGDHPCHGCWCVLHNTCTRKERCSEKPERFA-SE 554

455 AMTQDKVFLPVOECLSYPTCTOCRSQDDYCGWCVVEGCTKAECPRAEASHWLWSR 514

555 MKQCVRLT-VHPNNIS-VSQYNVLLVLETYNVPELSA--GVNCTFEDLSEMDGLVVGNOI 610

515 SKSCVAVTSAQPQNMSSRAQGEVQLTVSP--LPALSEEDLCLFGSPHPHARVEGEAV 572

611 QCYSPAKEVPRIITENG-DHHVVQLQKSKETGTMFASTSFVYNCVSINS-----CL 663

573 ICNSPSS--IP--VTPPGQDHVAVTIQLLRNGNIILTSYQYPPFYDCRQAMSEENLPCI 628

664 SCVESPYRCHWCKYRHVCTH-DPKTCSFQEG--RVKLPEDCPOLLRVDKILVPEVVKPI 720

629 SCVSNRWTQMDLRYHECREASNP---EDGIVRAHMEDESCPQFLGSPVLVPMNHTDV 685

721 TLKAKNLPOQSGQRGYECILNIQSGEQRVPA--LRFNSSSVQCNQTSYSYEGMEINN-- 776

Db 686 NFOGKNDL-----TVKSGSLHVGSDLLKFMFVPTWQESGTFAPRTPKLSHDA 732  
 QY 777 ---LPVELITVVMGHFNIDNPAQNVH--LYKCGAMRESGLCLKADDPFACGWCQPGQ 831  
 Db 733 NETLPLHLVYKSYGK-NID---SKLHVTLVYKSGFSGDCSLCRAANPDYRCACWGQSR 787  
 QY 832 CTLRQHCPAQESOWLELSGAKSCTNPRITEIPVTPGREGGTVKTRGENGLGEPDIA 891  
 Db 788 CVYEALC-----NTTSECPVPVITRIQETPLGGGIRITILGSLNGLVQAGDI- 835  
 QY 892 SHKVGAVECSPLVDGVIPEAOVICEMGEAKPSOHAGFVEICVAVCRPEF--MARSS--- 946  
 Db 836 QRISVAGNCSFQPERYSVSTRIVCVI-EAAETFTGGEVDV-----FGKLGRSPNV 888  
 QY 947 QLYYFMTLTSLDKSPSGMGSGTQVITITGNLGNAGSVVMFGKQPCCLFHRRSPSYVC 1006  
 Db 889 QFTFQPKPLS-VEPQOQPAAGGTTLTHGTHLDTGSOEDV-----RVTLNGVPC 937  
 QY 1007 NTSSSDEVELEKVSVOVDRAKI-----HQDLVFOYVEDPTVRIEPEWSIVSG 1054  
 Db 938 KVTGFGAQLCVQTPQATRGOMLLEVSYGSPVNPFGIFTTYRENVPVLAFAPEPLRSFASG 997  
 QY 1055 NTPIAVMGTHLDLQNPQIRAKHGKHEHINICEVLNATEMTCOA-----PALALGPDH--- 1107  
 Db 998 GRSINVTGQFSLIQRFAM-----VWIAEPLOSQWPPREAESLQPMVTVGTDYVFH 1048  
 QY 1108 -----QSOLTERPEFGFI-----LDNVOSLLILNKNTYYPNPVFEAFGPSGILEL 1155  
 Db 1049 NDTKVFLSPVAPPEPEYAINLTVLIEMDGHRALLTEAGAFVEYDPTFENP----- 1100  
 QY 1156 KPGPIILKGNLPPVAGGNKLVNTV-----LVGKPCPTV-TVSDVQLLCESEPNLIGR 1209  
 Db 1101 ---TGVVKVQVKNLHARGTNLANKWTLQEAFAFVGAERCTMTLTETDLYCEPEVQPP 1157  
 QY 1210 HK-----VMARVGMEYSFGMVYI---APDSPLSL-----PAIVSIAGV 1246  
 Db 1158 PKRQKQRTTNLPEFIVKFGSREWLGRVEYDTRVSDVPLSLILPLVIVPMVVVIAVS- 1216  
 QY 1247 GLLIIFIVANLIAYKRESDLTLKRLQOMDNLESVALKECAFELQTDIHELTS 1306  
 Db 1217 -----VCYWRKQQAEREYEKKSQLEGLSESVDRCKFEFTDMLMEMDQND 1266  
 QY 1307 LDGAGIPFLDYTYMVLV-----PGIEDHPVLRLDEYVQERQVEKGLKLPALINNK 1362  
 Db 1267 VHEAGIPVLVDYTYTDRVFFLPSKGDQDMITGKLDIPEPRPVVEQALYQFSNLLSK 1326  
 QY 1363 VFLSFIRLESQSFMSMDRGNVSLIMTVLOSLEYATDVLKQLLADLIDKNLESKNH 1422  
 Db 1327 SFLINFIHTLENQREFSARAKVVFASLLVALHGKLEYTYDINHTLFLLELLEQYVVVAKN- 1385  
 QY 1423 PKLLLRRTESVAEKMTNNFTLLYKFLKECAGEPLFLCAIKQOMKGPIDAITGEAR 1482  
 Db 1386 PKMLLRSETVVERMLSNWISICLYLQDSAGEPLYKLFKALKHQVEKGPVDAVQKAK 1445  
 QY 1483 YSLSSEDKLRQOQIDYKTLVLSVSPDNANSPEVPVKILNCDTITQVKEKILDAIFKNVPC 1542  
 Db 1446 YTLNDTGLGDDVEYAPLTVSVIVQDB-GVDAIPVVKVNLNCDTISQVKEKIIDQVYRQPC 1504  
 QY 1543 SHRPKAADMLEWRQSGARMILQDIEDITTKIENDMKRLNTLAHYQVPGDSVVVALVSKQV 1602  
 Db 1505 SCWPRPDSVLEWRPGSTAQ-ILSDLDLTSQREGRWKRNVNLMHYNVRDQATIL----- 1558  
 QY 1603 TAYNANNSTVTSASKEYENMIRYGS PDSLSRSTPMITPDLESYGVMWHLVKNHHEGD 1662  
 Db 1559 -----SKVGVSQPEDSQDQL---PGERHAL-----LEENRNVHLVRPTDEVD 1599  
 QY 1663 QKEGDRGS-----KMWSEIYITRLATKGTQKQFVDDLFEITFSTAHRGSLALPAIKY 1715  
 Db 1600 EGKSGKSVKEKTRKAITIYILTRLSVKGTLQOQFVDFNFQSVLAPGH---AVPPAPV 1656  
 QY 1716 MFDFLEQADKHGHDHVRHTWKSNCPLPLRFVWVMKIPQFVDFIDHKNSTIDACLSSVA 1775

Db 1657 FPDFLDEQAQKHNIQDEDTIHWNKNSLPLRFWNILKNPHFIEDVHVHVVDASLSVIA 1716  
 QY 1776 QTFWDCSTSEHRLGKDSNKLKYAKDIIPSYKNWVERYYSIDIGKMPAISQDDMMAYLAE 1835  
 Db 1717 QTFWDACTTRTEHKLSDSPSNKLKYAKEISTYKQWVEDYKGIQWQVSDQDMNTHLAE 1776  
 QY 1836 QSRMHMNEFTMGSALSEIFSVGVKYSSEILGPLDHDGCGKOKLAYKLEOV 1886  
 Db 1777 ISRAHTDSLNTLVALHQLYQYTKYDEIINALBEDPAAQKQLAFRLQOI 1827  
 RESULT 6  
 ABB11818  
 ID ABB11818 standard; peptide; 1849 AA.  
 XX  
 AC ABB11818;  
 DT 11-JAN-2002 (first entry)  
 XX  
 DE Human plexin-B1/SEP receptor homologue, SEQ ID NO:2188.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;  
 KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;  
 KW cytosstatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;  
 KW antifungal; vulerary; antulcer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157188-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US03800.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-457740/49.  
 DR N-PSDB; ABA09062.  
 XX  
 PT Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PT e.g. arthritis and cancer  
 XX  
 PS Claim 20; Page 259-260; 1963pp; English.  
 XX  
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
 CC invention also relates to vectors and recombinant host cells comprising a  
 CC nucleotide of the invention, methods of producing the novel polypeptides,  
 CC antibodies against the polypeptides, methods of detecting the nucleotides  
 CC or polypeptides in a sample, and methods of identifying compounds which  
 CC bind to polypeptides of the invention. Although novel, many of the  
 CC polypeptides of the invention have homology to known proteins, thereby  
 CC giving an insight into their probable biological activities, and hence  
 CC potential therapeutic applications. The polypeptides of the invention may  
 CC have various activities, including cytokine, cell proliferation or cell  
 CC differentiation activities; stem cell growth factor activity;  
 CC haematopoiesis regulatory activity; tissue growth activity;  
 CC immunomodulatory activity; activin- or inhibin-related activities;

chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be depending on oncogenesis, cancer cell proliferation or metastasis. CC  
 CC The invention on their biological activities, polypeptides and nucleotides of CC  
 CC conditions, e.g., by protein or gene therapy. Such conditions include CC  
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell CC  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis), CC  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease, CC  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal CC  
 CC vascular growth. Polypeptides involved with tissue regeneration and CC  
 CC repair (or nucleic acids encoding them) may be used to promote wound CC  
 CC healing (e.g., of burns, incisions and ulcers), while those with CC  
 CC immunomodulatory activities may be used in the treatment of viral, CC  
 CC bacterial and fungal infections in addition to immune disorders. CC  
 CC Polypeptides with growth factor activity may be used in cell cultures to CC  
 CC promote cell growth. For example, such polypeptides may be used to CC  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells CC  
 CC that can be used to augment or replace cells damaged by illness, CC  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides CC  
 CC may also be used in the diagnosis of the above conditions, and in drug CC  
 CC screening techniques. The present sequence represents a novel human CC  
 CC polypeptide of the invention. CC  
 CC

Sequence 1849 AA;

Query Match 25.7%; Score 2565; DB 22; Length 1849;

Best Local Similarity 33.1%; Pred. No. 3.2e-213; Matches 552; Conservative 343; Mismatches 716; Indels 260; Gaps 61;

24 TLLTRQAPLQKORSVYTFRGEPAAGNHLVVDERTGHIYLGAVNRIYKLSDLKLVLT 83  
 20 TLLGLLGGAGLSRPRKLDFFRSE--KELNHLAVDEASGVVYLGAVNALYQLDKQLQEQ 77  
 84 HETGDEPNKCYPRIVQTCNEPIITNNVNMKLLIDYKENRLIACGSLYOGICKLRL 143  
 78 VATGPVLNKKCTPIEASQCHE--AEITDNNVQLLVDPKRLVECGQLKGI CALRAL 136  
 144 ED--LFLKGEPRHKEHYLSGVNDSGVFVTVSVNSLDDKLFATAVDCKEYPTISS 201  
 137 SNISLRLFYEDSGSEKSFVANDGEGVATVGLVSVSTGPGGDRVLFVCKGNGPHDNGIIVST 196  
 202 RKLTKNSBADGMFAVHDEFVASMIKIPSDTFTIIPDPDIYVYVYFGSGNFVYFLTLP 261  
 197 RLLDRDTSREAFEAFTDRATYKAGYLSLNTQGF-----VAAFEDGVPVFFVFNQ 246  
 262 EMVSPGFTKEQVYTSKLVRLCKEDTAFNSVVEPIGCGSGVEYRLQAAVLSKAGAV 321  
 247 DK-HPARNRT-----LLARMCEDPNYSYLEMDLQCRDPDIH-----AAAF--GTC 290  
 322 LGRTGVHPDDLLFTVFSKGQKRMKSLDESALCIFILKQINDRIKRLQSCYRGEGTL 381  
 291 LAASVAAPSGSGLVAVFSRDS--RSSGPGAGLCLFPLDEVHAKMEANRACY--TGTR 346  
 382 DLAWLKVK----DIPCSSALLTIDNPF-CGLD-MNAPLGVSDVMRGIPVFTEDRDMTSV 435  
 347 EARDIFYKPFHGDICQGGHAPGSSKSPFCGSEHLPVPLGSRDGLRGATVLRGGGLNLTA 406  
 436 IAYVYNHSLAFVCTKSGKLLKIRVDGPRGNALQVETVOVDPGPVLRDMASKDHEQLY 495  
 407 TVAAENNHNTVAFGLSDGRILKYLIT-PDGTSSSEYDLSILVEINKRVKLDVLSGLSGLY 465  
 496 IMSERQLTRVPVSGCYQSCGCEGLSGDPHCGVLCVHNTCTRKERCERSKEPRAPA-SE 554  
 466 AMTQDKVFLPVOECLSTPTCTQCRDSQDPYCGWCVVEGRCRTRKAECPRAEASHLWSR 525  
 555 MKQCVRLT-VHPNNIS-VSQYNVLLVLETYNVPELSA--GVNCTFEDLSEMDGLVVGNOI 610  
 526 SKSCVAVTISAQPNMSRRAQGEVQLTVSP--LPALSEEDLCLFGESPFPARVEGEAV 583  
 611 QCYSPAAREVPRITENG--DHHVVOLOKSKETGMTFASTSFVFNCSVHNS-----CL 663  
 584 ICNPSSS--IP--VTPPGQDHVAVTIQLLLRGNIFLTSYQYFPFYDCRAMLEENLPCT 639

664 SCVESPYRCHWKYRHVCTH--DPKTCSEFORG--RVKLPEDCPOLLRRVDKILVPVEVIKPI 720  
 640 SCVSNRMTQWDLRYHCEASNP--EDGIVRAHMEDESCPQLGSPVLPVPMNHETDV 696  
 721 TLKAKMLPOQSGQRGYECILNTQSGEORVPA--LRPNSSSVQOCNTSYSEYGEIN-- 776  
 697 NFOGKLD-----TVKSSHLVGSDDLKFMPEVPTMQESGTFAFRTPKLSHDA 743  
 777 ---LPVELTVWNGHFNIDNPAQNVH--LYKCGAMRESGLCLKADDPACGWCQCPGQ 831  
 744 NETLPHLVKSYGK-NID-----SKLHVLYDCSFRSDCSLCRAANPDRACWCGQSR 798  
 832 CTLRQHCPAQESOWLELSGAKSCTNPRITEIIPVTPGREGTKVTIRGENLGLFRDIA 891  
 799 CYVEALC-----NTTSECPVITRIQETGPGGGIRITILGSLNGLVQAGDI- 846  
 892 SHVAVAGVCSPLVDGYIPAEQIVCEMGEAKPSQHAGFVIEICVACRPEF--MARSS--- 946  
 847 QRISVAGRNCSPQERYSVSTRIVCVI-EAAETPFTGGVEVDV-----FKLGRSPNV 899  
 947 QLYYFMTLTLSDLKPSRGPMSGGTQVITITGNLGNAGSVVVMFGKQPCLFHRRSPSIVC 1006  
 900 QTFQOPKPLS-VEPOGQPGAGGTTLITIHGTHLDTGSQEDV-----RVTLNGVPC 948  
 1007 NTTSSDEVLEMKVSVQVDRAKI-----HQDLVFPQYVEDPTIVRIEPEWSIVSG 1054  
 949 KVTKFGAQLQCVTGPQATRGOMLEVSYGSPVNPFGIIFTYRENPVLRFAPELRFSASG 1008  
 1055 NTPIAVWGTHLDLIQNPQIRAKHGKHEHINCEVLNATEMTCOA-----PALALGPDH--- 1107  
 1009 GRSINVTGQGSFLQIRFAM-----VVIAEPLOSQWQPPREAESLQPMVTGTVDFVH 1059  
 1108 -----QSDLTERPEEFGFI---LDNVQSLLILNKTNTYYPNPVEAFGSGILEL 1155  
 1060 NDTKVVLSPAVEPEEPAYNLTVLIEMDGHRAULRTEAGAFVYVDPDTFENF----- 1111  
 1156 KPQTPILKGNKILPPVAGNVKLVTV-----LVGEKPCTV-TVSDVQLLCESPNLIGR 1209  
 1112 ---TGGVKQVKNKIRARGTNLKNAMTLOEABAFVGAERCTMKTLTETDLYCEPEVQPP 1168  
 1210 HK-----VMARVGMESYFPMYI---APDSPLSL-----PAIVSIAVAG 1246  
 1169 PKRQRKRDTHNLPEFIVKFSREWVLGRVEYTRVSDVPLSLILPLVIVPMVVIVAS- 1227  
 1247 GLLIIFIVAVLITAYKESRESLTLKRLQMDNLSRVALECKEAEAELOTDIHELTS 1306  
 1228 -----VICYWRKSQAEREYKIKSQLEGBESVDRCKKEFTDLMIEMEDQTN 1277  
 1307 LDGAGIPFLDYRTYTMVLF---PGIEDHPVLRLDLEVPGYRQBERVEKGLKFAQLINNK 1362  
 1278 VHEAGIPVDYKYTYTDRVFLPSKQGDQVMTGKLDIPEPRPVVQALYQFSNLLNSK 1337  
 1363 VFLSIFRTLESQSPSMRDRGNVSLIMTVLOSLEYATDVLKQLLADLIDKNLESKNH 1422  
 1338 SFLINFIHTLENOPEFSARAKVYFASILLTVALHGKLEYTIDIMHTLFELELQVVAKN- 1396  
 1423 PKLLLRRTSVAEKMLTNMFTFLLYKFLKECAGEPLFLFCAIKQOQKEGPIDAITGEAR 1482  
 1397 PKMLRRTSETTVVERMUSNWMSCILYQYLDKDSAGEPLFKLFAKHQVEKGPVDVAKKAK 1456  
 1483 YLSLEDKLIRQIIDYKTLVLSCVSPDNANSPEVPVKILNCDDITQVKEKILDAIFKNVPC 1542  
 1457 YTLNDTGLLGDVVEYAPLTVSVIVQDB-GVDAIPVKVLNCDTISQVKEKIIDQVYRGPC 1515  
 1543 SHRPKAADMLEWRQSGARMILQDSDITTKIENDWKRNLTAHYQVDPDGSVALVSKQV 1602  
 1516 SCWPRPDSVVLEWRPGSTAQ-ILSDLDLTSQREGRKRVNLTLMHYNVRDGTATIL----- 1569  
 1603 TAYNAVNSTVSTASKEYENMIRYTGSPDLSRTPMITPDLESYKVMHVLKVNHEGD 1662  
 1570 -----SKGVSGQPPDSQODL---PGERHAL-----LEENRVHVLVRPTDEVD 1610

1663 QKEDGRGS-----KMSBIYLRLIATKGTLOKQFVDDLFETIFSTAHRSALPLAIKY 1715  
 1611 EGSKSGSVKEKERTKATIIYLITRLLSVKGTLQOQFVDFNFQSVLAPGH---AVPPAVKY 1667  
 1716 MFDFLDEQADKHGHDHPHVRHTWKSNCPLPLRFVWNNMIKNQFVDFDIHKNISITACLISVA 1775  
 1668 FDFLDEQAEKHNIQEDDTTHIKWNTSLPLRFVWNNILKNPHFIPDVHVHEVDSASLVIA 1727  
 1776 QTFWDCSTSEHRLGKDSNKLKYAKDIPSKNNWERYYSYDICKMPAISDDQDMNAYLAE 1835  
 1728 QTFMDACTRTEHKLSDSPSNKLLYAKEISTYKQWEDYTKGRQWQVQSDQDMNTHLAE 1787  
 1836 QSRMHNEFTMSALSIFSVGVKYSSEILGPLDHDQCGKQKLAYKLEQV 1886  
 1788 ISRAHTDSLNTLVALHQLQYQTKYDYDEIINALLEEDPAAQKWLAPRLQOI 1838

## RESULT 7

ABG21017  
 ID ABG21017 standard; Protein; 1870 AA.

XX AC ABG21017;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #21008.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC;

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS85204.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity

PS Claim 20; SEQ ID No 51376; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (II) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1870 AA;

Query Match 25.2%; Score 2521.5; DB 22; Length 1870;  
 Best Local Similarity 32.9%; Pred. No. 2e-209;  
 Matches 648; Conservative 342; Mismatches 721; Indels - 261; Gaps 62;

Qy 24 TLLTROAPILSOKORSFVTRGEPAGFNHVLVDERTGHIYLGAVNRVYKSSLDKLVLT 83

Db 40 TLLGLLGGASLRPRKLDFFRSE--KELNHLAVDEASGVVILGAVNALYQLDKAQLEQQ 97

Qy 84 HETGPDNDPKYPPRIVQTCNEPLTTTNNVNMKLLIDYKENRLIACSLYOGICKLRL 143

Db 98 VATGPALDNKKCTPIEASOCHE-AEMTDNVNPLLVDPKRLVECCGOLLKICALRAL 156

Qy 144 ED--LFLGEPYHKKEHYLGVNSESSEGVFVIVSYNLDLKLFIATAVDGKPEYFTISS 201

Db 157 SNISLRLFYEDSGSEKSFVASNDEGVATVGLVSTGPGGDRVLVFGKNGPHDNGIIVST 216

Qy 202 RKLTKNSEADGMPAYVFHDEFVASMIKIPSDTFTIIPDFIVVYVYVYVYVYVYVYVYV 261

Db 217 RLDRDTSREAFAYTDHATYKAGYLSNTQOF-----VAAFEDGPVVFVFNQ 266

Qy 262 EMVSPPGSTTKEQVYTKLVLCKEDTAFNSVVEVPIGCSRGVYRLLQAAVYLSKAGAV 321

Db 267 DK-HPARNRT-----LLARMCREDPNYSYLEMDLQCRDPDIH---AAAF---GTC 310

Qy 322 LGRTLGVHPDDLLFTVFSKQKRWKSLDESALCIFILKQINDRIKRLQSCYRGEGL 381

Db 311 LAASVAAPGSGRVLYAVFSRDS--RSSGGGAGLCLFLDLKVKHAKWEANRACY--TGTR 366

Qy 382 DLAWLKVK---DIPCSSALLTIDNF--CGLD--MNAPLGVSDMVRGIPVFTDRDRMTSV 435

Db 367 EARDIFYKPFHGDICCGHAPGSSKSPFCGSEHLFPVPLGSRDLRGTAQLVQGLNLTA 426

Qy 436 IAYVYKNSLAFVGTGSKLKKIRVDGPRGNALQYETVQVDPGVPGLRDMASKDEQLY 495

Db 427 TVAAENHNTVAFLGTSRILKLVLT-PDGTSSSEYDSILVEINKRVKDLVLSGDIGSLY 485

Qy 496 IMSERQLTRVPVSCQYQSGCEGLSGDPHCGWCVLHNTCTRKERCERSKEPRPA-SE 554

Db 486 AMTQDKVFLPVQECLSYPTCQCRDSQDPYCGWCVEGRCTKCAECPAEASHLWSR 545

Qy 555 MKQCVRLT-VHPNNIS-VSQYNVLVLVETYNVPELSA--GVNCTFEDLSEMDGLVVGNOI 610

Db 546 SKSCVAVTSAQFQNMRRRAQGEVQLTVSP--LPALSEDELLCLFGESPFPARVEAV 603

Qy 611 QCYSPAKEVPRITENG-DHHVVOLOLKSKEGTMTFASTSFVYNCVSHNS-----CL 663

Db 604 ICNPFSS--IP--VTPPGQDHVAVTIQLLRRGNIFLTSYQYFFYDCRAMSLEENLPCI 659

Qy 664 SCVESPYRCHWCKYRHVCTH-DPKTCSFOEG--RVKLPEDCPQLLRVDKILVPEVIKPI 720

Db 660 SCVSNRWTCQMDLRYHCREASPNP---EDGIVRAHMEDESCQFLGPSPLVPMNHETDV 716

Qy 721 TLKAKNLPQOSGQGRGYECILNIQSEQRVPA--LRFNSSSVQOCQNTSYSGEWMENN-- 776

Db 717 NFQGNLD-----TVKSSSLHVGSDLLKFMPEVPMQESGTFAPRTPKLSHDA 763

Qy 777 ---LPVELTVVWGHFNIDNPAQNVH--LYKCGAMRESCGLCKADPDPAFCWCGQPGO 831

Db 764 NETPLHLVYKSYGK-NID-----SKLHVLYNCSFGSDCSLCRAANPDYRCACWCGQSR 818

Qy 832 CTLRHQCPAQESQWLELSGAKSKCTNPRITEIPVTGREGGKTVTIRGENLGLEFRDIA 891

Db 819 CYVEALC-----NTTSECPPPVITRIQPETGPLGGGIRITILGSLNGLVQAGDI- 866

Qy 892 SHVKGVGCSPLDGYPABQIVCEMGEAKFSQAHGAFVEICVAVCRPEF--MARSS--- 946

Db 867 QRISVAGRNCSPQPERYSVSTRICVI-EAAETPFTGGVEVDV-----FGKLGSRSPNV 919





Db 1865 IAQAFIDACSIDLQKSDPTNKLKYAKIPEYRKIVQYRYKIQIDMTPLSEQEMNAHL 1924  
Qy 1834 AEQSRMHENFNTSALGEISYGVKYSSEILGLDHDQCGKQKLYAKLEOVITLM 1890  
Db 1925 ABESKYNQENFNTVMAEIIYKAKRYRQPMALLEANPTARTQLOKHFEQVVALM 1981

RESULT 9  
AAU00017  
ID AAU00017 standard; Protein; 1925 AA.

XX AC AAU00017;  
XX DT 09-MAY-2001 (first entry)  
XX DE Human Plexin-D1.

XX Human; Plexin-D1; semaphorin domain; hyperplasia; neoplasia; cancer;  
KW neurodegenerative disease; autoimmune disease; lupus; multiple sclerosis;  
KW inflammatory bowel disease; diabetes type I; rheumatoid arthritis;  
KW immunogen; antibody; MET related sequence; MRS.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH Peptide 1..42  
FT /label= Signal\_peptide  
FT Domain 1..1270  
FT /label= "Extracellular domain"  
FT Protein 43..1925  
FT /label= "Mature Plexin-D1"  
FT Domain 43..600  
FT /label= "Semaphorin domain"  
FT Region 541..620  
FT /label= "Primary Met related sequence, MRS"  
FT Region 691..770  
FT /label= "Secondary MET related sequence, MRS"  
FT Region 831..910  
FT /label= "Tertiary MET related sequence, MRS"  
FT Region 900..1110  
FT /label= "G-P repeat region"  
FT Domain 1270..1293  
FT /label= "Transmembrane domain"

XX WO200114420-A2.  
XX PD 01-MAR-2001.  
XX PF 25-AUG-2000; 2000WO-US23365.  
XX PR 25-AUG-1999; 99US-0150576.  
XX PA (UYTO-) UNIV TORINO.  
XX PA (REGC) UNIV CALIFORNIA.

XX Artigiani S, Comoglio PM, Goodman CS, Tesler-Lavigne M;  
XX Tamagnone L;  
XX WPI; 2001-226610/23.  
XX N-PSDB; AAS00022.

XX New plexin polynucleotides and polypeptides, useful in diagnosis,  
XX therapy and in producing compounds for treating diseases involving  
XX aberrant cell growth (e.g. cancer) or immune regulation (e.g.  
XX autoimmune diseases)

XX Claim 4; Page 64-69; 79pp; English.

XX The sequence represents Human Plexin-D1. Plexins are large transmembrane  
XX proteins whose extracellular domain shares homology with Scatter factor  
XX receptors and contain an approximately 500 amino acid Semaphorin domain.  
XX The plexin polynucleotides and polypeptides, and plexin-specific binding  
XX agents are useful in diagnosis, therapy and in the biopharmaceutical

CC industry. In particular, the plexin polynucleotides and polypeptides are  
CC useful for generating compounds (e.g. plexin-specific binding agents or  
CC antibodies) for treating or diagnosing a disease or disorder involving  
CC aberrant cell growth (e.g. hyperplasia, neoplasia, cancer or  
CC neurodegenerative disease), or diseases or disorders involving aberrant  
CC immune regulation (e.g. autoimmune diseases such as lupus, inflammatory  
CC bowel disease or diabetes type I), or immunosuppressive diseases such as  
CC multiple sclerosis or rheumatoid arthritis.  
XX  
SQ Sequence 1925 AA;

Query Match 24.7%; Score 2463.5; DB 22; Length 1925;  
Best Local Similarity 31.8%; Pred. No. 2.4e-204;  
Matches 644; Conservative 338; Mismatches 745; Indels 301; Gaps 60;

Qy 15 LLMVG---MGSSTLLTRQAPLSQKORSFVTRGEPAEGFHLVVDERTGHIVGAVNRI 71  
Db 36 LLLGGAARAGALEIQRFPPTPT-----NNFALDGAAGTVILAAVNRL 79  
Qy 72 YKLS-SDLKVLVTHETGPDENPKYPPRIQV-TCNEPLTTNNVNKMLLDYKFNRLIA 129  
Db 80 YQLSGANLSLEAEAAVGPVDPSPCHAPQLPQASCEHPRLTDYNNKILQLDPCQGLVVV 139  
Qy 130 CGSLYQGI CKLRLLEDL FKLGEPIYHKEHYLSGVNBSGSGVFGIVSYNSLD----- 180  
Db 140 CGSIYQGFQQLRRRGNISAVAVRFPAPPAEPVTPPSPMLNVAANHFNASTVGLVLPPA 199  
Qy 181 -----DKLFATVADGKPEYF-----PTISSRLTKNSEADGMFAYVPHD 220  
Db 200 AGAGSRLLVGATYTYGSGSFFPNRSLDRHRENTPEIAIRSLDTRGDIAKLFTFDLN- 258  
Qy 221 EFVAMIKIPSDTFTIIPDFIYVYVFGSGNFVYFLT--LQPEMVSPPGSTT----- 271  
Db 259 -----PSD-----DNILIKQGAKEQHLKGFVSAFLHPS-DPPGQAQSYAYLALN 302  
Qy 272 -----KEQVYTSKLVRLCKEDTA-----FNSVVEVPIGERSGVEYRLLQAAYLSK 317  
Db 303 SEARAGDKESQARSLLARICLPHGAGGDGAKKLTESYIQLG-----LQCGAGGAG 350  
Qy 318 AGAVLGRITLGVHPDDLLFTVFSKGOKRKMKSLSDESALCIFIILKQINDRIKERLQSCY-- 375  
Db 351 RGDLYSRLVSVFPARERLFVAFERPOGSPAARAAPALCAPFADVRAARAARTACFVE 410  
Qy 376 -----RGEGB-----TDLAWLWKVDI PCSSALLTIDDNFCGLDMNAPLGVSD 417  
Db 411 PAPDVAVLDSVVQGTGPACERKLNIQ-LQPEQLDCGA-----HLQPLSLTLQ 458  
Qy 418 MVRGIPVFTEDRDMTSVIAYVYKHNHSLAFVGTGSKGLKIRVDGPRGNALQYETVQVVD 477  
Db 459 PLKATPVFR--APGLTSVAVASVNNYTVAPLGTVNGRLKLN-----ESMQVVS 507  
Qy 478 -----PGPVLDMAFS-KDHEQLYIMSERQLTRVPVSESCGYQSCGECGLSGDPHCG 528  
Db 508 RRVTVAYGEPVHHVMQDFPADSGYLYLMTSHQMARVKVAACNVHSTCGDCVGAADAYCG 567  
Qy 529 WCVLHNTCTRKERCERSKEPRRFASEMK---QCVRLLTVPNNISVSQ-YNVLVLETYNV 584  
Db 568 WCALETRCTLOQDCTNSSQOHFWTSASEGSRCPAMTVLPSEIDVRQEPYPMIQLSGSL 627  
Qy 585 PELSA-GVNCTF-----EDLSEMDGLVVGNOIQCVSPAAKEVPRIITENGHHVVOLQLKS 639  
Db 628 PSLSGMEMACDYGNNIRTVARVPGAPGHQIAYCNLLPRDQFPFPFPNQDH--VTVMENV 685  
Qy 640 IETGTMFASTSFVFNCS-----VHNSCLSCVESPYRCHWKYRHVCTHDPKTCSEQEG 693  
Db 686 RVNGRNIVKANFTIYDCSRTAQVYPHTACTSCLSAQWPCFWCSQOHCSCVSNQSCASPN 745  
Qy 694 RVKLPDPCOLLRVDKILVPEVIKPIITLAKNLPQPSQORGYECILINIGSQRVPAL 753  
Db 746 PTS-PODCPTLLSPAPVPTGSGSNILVPLANTAFQGA--ALECSF---GLBEIFEAV 799  
Qy 754 RFNSSSVQC-QNTSYSYEGMEINNLVPELVVWNGHFHNIDNPQAKVHLYKCGAMRSCG 812

800 WYNESVVRCDQVVLHTRRSQV--FPLSLQLKGRPARFLDSPERMTVMVYNCAMGSPDCS 857  
813 LCL-KADDPFACWCOGPGOCTLRQHCAPAGESOWLESLGAKSKCTNPRITETIIPVTGPRE 871  
858 QCLGREDLGLHCWSDG---CRLRG--PLQ-----PWAGT---CPAPEIRAIPLSLGPLD 904  
872 GGTQVTRIGENLGLPRDIAHVKVAGVECSPLVDGYIPAEQIVCEMGEAKPSOHAGFVE 931  
905 GGTLLTIRGNLGRRLSDVAGWVGIVGACEPLPDRYTVSEEVICVTGPA-PGPLSGVVT 963  
932 ICVAVCRPEPMARSSQLYFYMTLTLSDKPSRPMSCGTQVITGTNLNAGSNVVMFG- 990  
964 V-----NASKEGSRDRFSVVLPLVHSLPTMGPKAGGTRITIHGNDLHVSGELQVLND 1018  
991 KQCLFHRRSPSVIVCTTSSDEVLKMSVQV---DRAKHQDLVFOYVEDPTIVRIEP 1047  
1019 TDFCTELMRTDTSIAC--TWPEGALPAPVPCVFRERRGCVGHGMLTFTWQNPVITAI 1076  
1048 EWSIVSGNTPIAVMGTHLDLIQNPIRAKHGKKEHINICEVLNATEMTCOAPALALGPDH 1107  
1077 RRPVSGGRITVAGERFHMVQVSMVHHIGREP-TLCKVLNSTLITCPSGALSNSA 1135  
1108 QSDLTERPEEFGFLDNVQSLIL-----NKNFTYYPNPVPEAFSPSGILELKPCTPI 1161  
1136 PVDFFINGRAVADVAEVAEELLDEEAQRGRFRFLDYLNPQFSTAKREKWKIKHHPCEPL 1195  
1162 IL---KKNLIPPVAGNVKLVNLYLGEKPCPTVSDVOLLCEPN-----LIGRHKVM 1213  
1196 TLVIHKEODSL-----GLQSHYRKVGQVSCDIQIVSDRIIHCVSNEISGAAGQOLPIT 1250  
1214 ARVGMEYSPGMVYIAPDSPLSLPAIVSIYAVAGLLIIFIVAVLIYAKRKRSRSDLTMR 1273  
1251 IQGVNFQNTIATLQGS---ETAIIVSIVICSVLLLSVVA-LFVFCCKSRRAERYWQK 1306  
1274 LQOMQNLERSVALECKEAPAELOTDIHLTSDLD-GAGIFLDYRTYTMVRLP----- 1327  
1307 TLQMBEMESQIREIRKGAELQTDLTDLKELNSQIIPFLEYKHFTVTFPPKCSL 1366  
1328 -----GTEDHPVRLDLVPVGRQERVEKGLKPAQLINNKVFLLSF 1368  
1367 YEERYVLPSTLNSQSGSQAOETHPLLGWKIPESCRPMEEGISVFSLLNNKHFLIVF 1426  
1369 IRTLESORFSMRDRGNVASLIMTVLOSQKRYATDVLKOLLADLIDKNLESKNHPKLLLR 1428  
1427 VHALEQQKDFAVDRCSLASLTLTIALHGKLEYTTSIMKELLVDLIDAS--AAKNPKMLR 1484  
1429 RTESVAEOMLTNFTFLYKFLKCEGAPSLFCAIKQOMEKGPIDAITGEARYSLSD 1488  
1485 RTESVVERKMTNWSICMYSCSLRETGVGFEFFLLCAIKQIQKNGSIDAITGARTLNEE 1544  
1489 KLIRQIDYKTLVLSCVSPDNANSPEVPVKILNCDTITQVKEKILDAIFKNVPCSHRPKA 1548  
1545 WLLRENIKAPNLN-VSQFGCMDSLSVRAMDTTLQVKEKILEAFCKNVVPSQWPA 1603  
1549 ADMLEWROGSGARMLODEDITTKIENDWKRNLTAHVQVPDGSVA--LVSKQVATYN 1606  
1604 EDVLEWFASSSTQVILRDLDTSVVEDGRKKLNTLAHYKIPEGASLAMSLLDK----- 1658  
1607 AVNSTVSTASKYENMIRYTGSPDLSRPTMITPDLESQVKKWHLV-KNHEHGDOKE 1665  
1659 ---DNTLGRVK-----DLDT-EKYFHLVPLTDELAEPKK 1688  
1666 GDRGS---KMWSEIYLTRLLATKGLQKFVDDLFTFSTAHRSALPLAIKIMFDFLDE 1722  
1689 SHRQSHRKVLPPIYLTLLSTKGLTQKFLDLDLFRALISI--REDKPLAVKYFFDFLE 1746  
1723 QADKHGHHDPVHRTWKSNCPLRFRWNMKPQVFDTHKNSITDACLVSUAQTFMDSQ 1782  
1747 QAEKRGISDPDTHIWKTNLSPLRFWNLKPNQFVFDIDKTDHIDACLVSUAQAFIDAC 1806  
1783 STSEHLGKDSNKLKYAKDIPSYKNWERYSDIGKMPAISQDQMNAYLAEQSRMHN 1842  
1807 SISDLQGLKDSPTNLLIYAKEIPEYRKIVQRYKQIQDWTPLSEQEMNAHLAESRYQN 1866

QY 1843 EFTNMSALSIFSVGKYSEIILGPLDHDQCGKQKLAYKLEQVITLM 1890  
Db 1867 EFTNVAMAEIFRSPKRYQIQMAALEANPTARTLOHQKEQVVALM 1914  
RESULT 10  
AAU00016  
ID AAU00016 standard; Protein; 1788 AA.  
XX AAU00016;  
AC AAU00016;  
XX 09-MAY-2001 (first entry)  
XX Human Plexin-B3.  
XX Human; Plexin-B3; semaphorin domain; hyperplasia; neoplasia; cancer;  
KW neurodegenerative disease; autoimmune disease; lupus; multiple sclerosis;  
KW inflammatory bowel disease; diabetes type I; rheumatoid arthritis;  
KW immunogen; antibody; MET related sequence; MRS.  
XX Homo sapiens.  
XX Location/Qualifiers  
FT Key 1..23  
FT Peptide /label= Signal peptide  
FT Domain 1..1099  
FT FT /label= "Extracellular domain"  
FT Protein 24..1788  
FT FT /label= "Mature Plexin-B3"  
FT Domain 24..507  
FT FT /label= "Semaphorin domain"  
FT Domain 1100..1119  
FT FT /label= "Transmembrane domain"  
XX WO200114420-A2.  
XX 01-MAR-2001.  
XX 25-AUG-2000; 2000WO-US23365.  
XX 25-AUG-1999; 99US-0150576.  
XX (UYTO-) UNIV TORINO.  
XX (REGC) UNIV CALIFORNIA.  
XX Artigiani S, Comoglio PM, Goodman CS, Tesier-Lavigne M;  
PI Tamagnone L;  
XX WPI; 2001-226610/23.  
DR N-PSDB; AAS00021.  
XX New plexin polynucleotides and polypeptides, useful in diagnosis,  
PT therapy and in producing compounds for treating diseases involving  
PT aberrant cell growth (e.g. cancer) or immune regulation (e.g.  
PT autoimmune diseases)  
XX Claim 4; Page 58-62; 79pp; English.  
XX The sequence represents Human Plexin-B3. Plexins are large transmembrane  
XX proteins whose extracellular domain shares homology with Scatter factor  
XX receptors and contain an approximately 500 amino acid semaphorin domain.  
XX The plexin polynucleotides and polypeptides, and plexin-specific binding  
XX agents are useful in diagnosis, therapy and in the pharmaceutical  
XX industry. In particular, the plexin polynucleotides and polypeptides are  
XX useful for generating compounds (e.g. plexin-specific binding agents or  
XX antibodies) for treating or diagnosing a disease or disorder involving  
XX aberrant cell growth (e.g. hyperplasia, neoplasia, cancer or  
XX neurodegenerative disease), or diseases or disorders involving aberrant  
XX immune regulation (e.g. autoimmune diseases such as lupus, inflammatory  
XX bowel disease or Diabetes Type I), or immunosuppressive diseases such as  
XX multiple sclerosis or rheumatoid arthritis.

[illegible]

QY 1877 OKLAYKLEQVITLM 1890  
DB 1768 LQLACLQVAAVLV 1781  
RESULT 11  
ID ABB11753  
XX ABB11753 standard; peptide; 362 AA.  
AC ABB11753;  
XX  
DT 11-JAN-2002 (first entry)  
XX  
DE Human NOV/plexin-A1 homologue, SEQ ID NO:2123.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; anti-inflammatory;  
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
KW cytosatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
KW antifungal; vulnery; antiulcer.  
OS Homo sapiens.  
OS  
PN WO200157188-A2.  
PN  
PD 09-AUG-2001.  
PD  
XX  
PF 05-FEB-2001; 2001WO-US03800.  
PF  
XX  
PR 03-FEB-2000; 2000US-0496914.  
PR  
PR 27-FEB-2000; 2000US-0560875.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Drmanac RT;  
PI  
XX WPI: 2001-457740/49.  
DR  
DR N-PSDB; ABA08997.  
XX  
PT Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
PT e.g. arthritis and cancer -  
XX  
PS Claim 20; Page 239; 1963pp; English.  
XX  
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
XX invention also relates to vectors and recombinant host cells comprising a  
XX nucleotide of the invention, methods of producing the novel polypeptides,  
XX antibodies against the polypeptides, methods of detecting the nucleotides  
XX or polypeptides in a sample, and methods of identifying compounds which  
XX bind to polypeptides of the invention. Although novel, many of the  
XX polypeptides of the invention have homology to known proteins, thereby  
XX giving an insight into their probable biological activities, and hence  
XX potential therapeutic applications. The polypeptides of the invention may  
XX have various activities, including cytokine, cell proliferation or cell  
XX differentiation activities; stem cell growth factor activity;  
XX haematopoiesis regulatory activity; tissue growth activity;  
XX immunomodulatory activity; activin- or inhibin-related activities;  
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or  
XX thrombolytic activities; receptor or ligand activities; or may be  
XX involved in oncogenesis, cancer cell proliferation or metastasis.  
XX Depending on their biological activities, polypeptides and nucleotides of  
XX the invention are useful for preventing, treating or ameliorating medical  
XX conditions, e.g., by protein or gene therapy. Such conditions include

CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a novel human  
CC polypeptide of the invention.  
XX  
SQ Sequence 362 AA;  
Query Match 15.8%; Score 1573.5; DB 22; Length 362;  
Best Local Similarity 84.5%; Pred. No. 5.5e-128;  
Matches 300; Conservative 22; Mismatches 32; Indels 1; Gaps 1;  
QY 1420 KNHPKLLLRRTESVAEKMLTNWFTFLYKFLKESAGEPLFSLFCAIKQOMEKGPIDAITG 1479  
DB 9 KNHPKLLLRRTESVAEKMLTNWFTFLYKFLKESAGEPLFSLFCAIKQOMEKGPIDAITG 68  
QY 1480 EARYSLSEDKLIRQIDYKTLVLSCVSPDNANSPVPVKILNCDTITOVKEKILDAIFKN 1539  
DB 69 EARYSLSEDKLIRHLDYKTLVLCVNPENENAPVPVKGDLCDTGTQAKELLDAAKYG 128  
QY 1540 VPCSHRPKADMDLEWRQGGARMILQDEITTKIENDWKRLNTLAHYQVPGSVVALVS 1599  
DB 129 VPYSORPKADMDLEWRQGRMARIIQDEVTTKIDNDWKRLNTLAHYQVTDGSSVALVP 188  
QY 1600 KQVTAYNAVNNSTVSRTSASKVENIRYTGSPDSLSRSTPMITPDLESQVKMHLVKNHE 1659  
DB 189 KQTSAYNISNSTFTK-SLSRYESMLRTASSPDSLSRSTPMITPDLESQVKMHLVKNHD 247  
QY 1660 HGDOKEGDRGSKMVSEIYLTRLLATKGTLOKFVDDLFEFIFSTAHRSALPLAIKYMDF 1719  
DB 248 HLDQREGDRGSKMVSEIYLTRLLATKGTLOKFVDDLFEFIFSTAHRSALPLAIKYMDF 307  
QY 1720 LDEQADKHGIHDHPVHRHTWKSNCPLRFLFWNNMKNPQFVFDIHKNSITDACLSSV 1774  
DB 308 LDEQADKHQIHHDADVHRHTWKSNCPLRFLFWNNVKNPQFVFDIHKNSITDACLSSV 362  
RESULT 12  
AAM93421  
ID AAM93421 standard; Protein; 716 AA.  
XX  
AC AAM93421;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide, SEQ ID NO: 3042.  
XX  
KW Human; full length cDNA; cDNA synthesis; oligo-capping.  
XX  
OS Homo sapiens.  
XX  
PN EP1130094-A2.  
XX  
PD 05-SEP-2001.  
XX  
XX 07-JUL-2000; 2000EP-0114089.  
PF  
XX 08-JUL-1999; 99JP-0194486.  
PR  
PR 11-JAN-2000; 2000JP-0118774.  
PR  
PR 02-MAY-2000; 2000JP-0183765.  
XX

(HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
WPI; 2001-524255/58.  
N-PSDB; AAK94342.

830 Primers useful for synthesizing full length cDNA clones and their  
use in genetic manipulation -

Disclosure; SEQ ID NO 3042; 1380pp + sequence listing; English.

The invention relates to primers for synthesizing full length cDNA  
clones. 830 cDNA molecules encoding a human protein have been  
isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
molecules have been determined. Primers for synthesizing the full length  
cDNA are useful for clarifying the function of the protein encoded by  
the cDNA. The full length clones were obtained by construction of full  
length enriched cDNA libraries that were synthesised by the oligo-capping  
method. The primers enable the production of the full length cDNA easily  
without any special methods. The present sequence is a polypeptide  
provided in the specification.

Note: The sequence data for this patent did not form part of the printed  
specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 716 AA;

Query Match 14.9%; Score 1486; DB 22; Length 716;  
Best Local Similarity 41.9%; Pred. No. 8.1e-120;  
Matches 309; Conservative 145; Mismatches 207; Indels 76; Gaps 14;

QY 1185 VGEKPTV-TVSDVQLCLSPNLIGRHK-----VMARVGMEYSFGWVI-- 1228  
DB 10 VCAERTMTLTETDLYCPEPPVPPKQKRDTHNLPDEFIVKGSREWVLRVEYDT 69

QY 1229 -APDSPLSL-----PAIVSTAVAGGLLIIFIVAVLIAYKRSRSDTLKRLQMDN 1280  
DB 70 RVSVDPLSLILPLVIVPMVWVIAVS-----VYCWKRSQAERYEYKIQLEG 118

QY 1281 LBSRVALECKEAFAELOTDIHELTSLDGAGIPFLDYRTYTRVLP-----PGIEDHPVLR 1336  
DB 119 LEESVRDRCKGFTDLIMEQDQTVHAGIPVLDYKTYTDRVFLPSKDGDKVMITG 178

QY 1337 DLEVPGYRVERVEKGLKFAQLINNKFLLSFTLRLESQPSMRDRGNVASLMTVLQS 1396  
DB 179 KLDIPEPRPVVEQALYQFNSLLNSKFLNFHTLENQEFSAKAVYFASLLTVALHG 238

QY 1397 KLEYATDLVKQLLADLIDKNSKNHPKLLLRRTESVAEKMLTNWFTFLYKFLKSCAGE 1456  
DB 239 KLEYTYDTHMTLFLLELQVWVAKN-PKMLRSETWVERMLSNWMSICLYQLKDSAGE 297

QY 1457 PLFSLFCAIKQOMEKPIDAITGEARYSISEDKLIHQQIDYKTLVLSCVSPDNANSPEVP 1516  
DB 298 PLYKLFKAIHQVEKGPDAVQKAKYTLNDTGLLGDDEYAPLTYSVIVQDE-GVDAIP 356

QY 1517 VKILNCDDTTQVEKILDAIFKNVPSCHPKAADMLEWROGSGARMILQDEDTTKIEN 1576  
DB 357 VKVLNCDDTTSQVEKILIDYRGQPCSCWRPPSVVLEWPGSTAQ-ILSDLDTSQREG 415

QY 1577 DWKRLNLTALHYQVPGDSVVALVSKQVYAYNAVNNSTVSRTSASKYENMIRYTGSPDSLRS 1636  
DB 416 RWKRVNTLMHYNVRCGATLIL-----SNVGSQQPEDSQDQL---PGERHAL-- 459

QY 1637 RTPMTPTDLESGVWVWHLVKNHGHGQKEDGRGS-----KMSIYILTRLLATGTLQ 1689  
DB 460 -----LEENRVWHLVRPTDEVDGSKSGKSVKEKRTKAITRIYTLRLSVRGTLQ 511

QY 1690 KFVDDLPTFTFSTAHGSGALPLAIKVMDFDLSDQADKHGHDHPHVRHTWKNCLPLREFW 1749  
DB 512 QFVDFNFRSVLAPGH---AVPRAVKTFDFDLSDQAEKHNIQDEDTIHWKNTSLPLRFW 568

QY 1750 NMIKNPQFVDIHKNSITDACLSVAQAQTFMDS CSTSEHRLGKDSPSNKLIAKYADIPSYKN 1809

PA (HELI-) HELIX RES INST.  
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX WPI; 2001-524255/58.  
DR N-PSDB; AAK94342.  
XX 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation -  
XX Disclosure; SEQ ID NO 3042; 1380pp + sequence listing; English.  
PS The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been  
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
CC molecules have been determined. Primers for synthesizing the full length  
CC cDNA are useful for clarifying the function of the protein encoded by  
CC the cDNA. The full length clones were obtained by construction of full  
CC length enriched cDNA libraries that were synthesised by the oligo-capping  
CC method. The primers enable the production of the full length cDNA easily  
CC without any special methods. The present sequence is a polypeptide  
CC provided in the specification.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.  
XX XX Sequence 716 AA;

RESULT 13  
AAM93435  
ID AAM93435 standard; Protein; 484 AA.  
XX  
AC AAM93435;  
XX  
DT 06-NOV-2031 (first entry)  
XX Human polypeptide, SEQ ID NO: 3070.  
DE Human; full length cDNA; cDNA synthesis; oligo-capping.  
XX  
KW Homo sapiens.  
XX  
OS EPI130094-A2.  
XX  
PN 05-SEP-2001.  
XX  
PD 07-JUL-2000; 2000EP-0114089.  
XX  
PF 08-JUL-1999; 99JP-0194486.  
PR 11-JAN-2000; 2000JP-0118774.  
PR 02-MAY-2000; 2000JP-0183785.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX  
DR WPI; 2001-524255/58.  
DR N-PSDB; AAK94356.  
XX  
XX 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation -  
XX  
PS Claim 8; SEQ ID NO 3070; 1380pp + sequence listing; English.  
XX  
CC The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been  
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
CC molecules have been determined. Primers for synthesizing the full length  
CC cDNA are useful for clarifying the function of the protein encoded by  
CC the cDNA. The full length clones were obtained by construction of full  
CC length enriched cDNA libraries that were synthesised by the oligo-capping  
CC method. The primers enable the production of the full length cDNA easily  
CC without any special methods. The present sequence is a polypeptide  
CC encoded by a full length human cDNA of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.  
XX  
SQ Sequence 484 AA;

Query Match 13.8%; Score 1376; DB 22; Length 484;  
Best Local Similarity 55.3%; Pred. No. 1.6e-110;  
Matches 267; Conservative 70; Mismatches 106; Indels 40; Gaps 5;

QY 23 STLLTRQAPLSQKOR-----SFTYRGEPAE- 49  
DB 6 SRLLT--AAPLSMEQRPMWPALEVDSRVLLSVVWVLLAPPAGMPQFSTFHSNEDW 63  
QY 50 GFNHLVVDERTGHIYLGAVNRIVKLSDDLKVLVTHETGPDENPKCYPPIVQTCNEPLT 109

569 NILKNPHFIEDVHVHVVDASLSVIAQTFMDACTRTEHKLSDSPSNKLLYAKEISTYK 628  
1810 WVERYSIDGKMPAISQDMMNAYLABOSRMMNFNTMSALSEIFSYGVKYSSEILGPLD 1869  
629 MVEDYKGIQRMQVQSDQDMNTHLAEISRAHTDSLNTLVALHQLYQYTKYDEIINALE 688  
1870 HDOQCGKOKLAYKLEQV 1886  
689 EDPAAQKQQLAFRLQOI 705

Db 64 T F N H L T V H O G T G A V Y V G A I N R V K L T G N L I Q V A H K T G P E E D N K S C Y P P L I V O P C S E V L T 123  
QY 110 T T N N V K M L L I D Y K E N R L I A C G S I Y Q G I C K L R L E D L F K L G P V H K K E H Y L S G W N E S G V 169  
Db 124 L T N N V K L L I D Y S E N R L I A C G S I Y Q G V C K L R L D D L F I L V E P S H K K E H Y L S G V N K T G T M 183  
QY 170 F G V I S V N L D D L K F T A T A V D G R P E Y P T I S S R K L T K N S E A D G M F A Y V F H D E F V A S M I K I 229  
Db 184 Y G V I V S E G E D K L F I G T A V D G Q D Y F P T L S S R K L P E D P E S S A M L D Y E L H S D F V S S L I K I 243  
QY 230 P S D T F I I P D I Y Y V G F S G M F V P L T I O P E N V S P P G --- S T T K E Q V T S K L V R L C K E 286  
Db 244 P S D T A L V S H F D I F Y I G F A S G G F V Y L T V Q P E -- T P E G V A I N S A G D L F Y T S R I V R L C K D 301  
QY 287 D T A F N S V E P I C E R S G V E R L L O A Y L S K A G A V L G R T L G V H P D D L L F T V S K G O K R K 346  
Db 302 D P K F H S V S L P F G C T R A G V E R L L O A Y L A P G S L A Q A F N I T S Q D D V L F A I F S K G O Q Y 361  
QY 347 M K S L D E S A L C I F I L K Q I N D R I K E R L Q S C Y R G E G T L D L A W L K V D I P C S S A L L I T I D D N F C G 406  
Db 362 H H P P D S A L C A F I R A I N L O I K G R L Q S C Y G E G N L E L N W L L G K D V Q C T K A P V P I D D N F C G 421  
QY 407 L D W N A P L G S D M V R G I P V F T E D R D M T S V I A Y V Y K N S L A P V G T K S G K L K I R V D G S R G N 466  
Db 422 L D I N Q P L G S T P V E G L T Y T T S R D R M T S V A S Y V Y G V F V G T K S G K L K I R A D G P P H G 481  
QY 467 A L Q 469  
Db 482 G V Q 484

RESULT 14  
AAV13373  
ID AAV13373 standard; Protein; 552 AA.

XX AC AAY13373;  
XX DT 25-JUN-1999 (first entry)  
XX DE Amino acid sequence of protein PRO235.  
XX KW Secreted protein; transmembrane protein; human; enterocolitis;  
KW Zöllinger-Ellison syndrome; gastrointestinal ulceration;  
KW congenital microvillus atrophy; skin disease; cell growth;  
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;  
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;  
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophla areata;  
KW anti-thrombotic; wound healing; tissue repair.  
XX OS Homo sapiens.  
XX PN WO9914328-A2.  
XX PD 25-MAR-1999.  
XX PF 16-SEP-1998; 98WO-US19330.  
XX PR 25-NOV-1997; 97US-0066840.  
PR 17-SEP-1997; 97US-0059113.  
PR 17-SEP-1997; 97US-0059115.  
PR 17-SEP-1997; 97US-0059117.  
PR 17-SEP-1997; 97US-0059119.  
PR 17-SEP-1997; 97US-0059121.  
PR 17-SEP-1997; 97US-0059122.  
PR 17-SEP-1997; 97US-0059184.  
PR 18-SEP-1997; 97US-0059263.  
PR 18-SEP-1997; 97US-0059266.  
PR 15-OCT-1997; 97US-0062125.  
PR 17-OCT-1997; 97US-0062285.  
PR 17-OCT-1997; 97US-0062287.  
PR 21-OCT-1997; 97US-0063486.  
PR 24-OCT-1997; 97US-0062814.

PR 24-OCT-1997; 97US-0062816.  
PR 24-OCT-1997; 97US-0063045.  
PR 24-OCT-1997; 97US-0063120.  
PR 24-OCT-1997; 97US-0063121.  
PR 24-OCT-1997; 97US-0063127.  
PR 24-OCT-1997; 97US-0063128.  
PR 27-OCT-1997; 97US-0063329.  
PR 27-OCT-1997; 97US-0063327.  
PR 28-OCT-1997; 97US-0063541.  
PR 28-OCT-1997; 97US-0063542.  
PR 28-OCT-1997; 97US-0063544.  
PR 28-OCT-1997; 97US-0063549.  
PR 28-OCT-1997; 97US-0063550.  
PR 28-OCT-1997; 97US-0063564.  
PR 29-OCT-1997; 97US-0063435.  
PR 29-OCT-1997; 97US-0063704.  
PR 29-OCT-1997; 97US-0063732.  
PR 29-OCT-1997; 97US-0063738.  
PR 29-OCT-1997; 97US-0063734.  
PR 29-OCT-1997; 97US-0064215.  
PR 29-OCT-1997; 97US-0063735.  
PR 31-OCT-1997; 97US-0063870.  
PR 31-OCT-1997; 97US-0064103.  
PR 03-NOV-1997; 97US-0064248.  
PR 07-NOV-1997; 97US-0064809.  
PR 12-NOV-1997; 97US-0065186.  
PR 17-NOV-1997; 97US-0065846.  
PR 18-NOV-1997; 97US-0065693.  
PR 21-NOV-1997; 97US-0066120.  
PR 21-NOV-1997; 97US-0066364.  
PR 24-NOV-1997; 97US-0066772.  
PR 24-NOV-1997; 97US-0066466.  
PR 24-NOV-1997; 97US-0066770.  
PR 24-NOV-1997; 97US-0066511.  
PR 24-NOV-1997; 97US-0066453.  
XX  
FA (GETH ) GENENTECH INC.

Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;

WPI; 1999-229533/19.  
N-PSDB; AAX52244.

New isolated human genes and polypeptides used in, e.g. treatment of gastrointestinal ulceration

Claim 12; Fig 62; 320pp; English.

AAV13344-403 represent secreted and transmembrane human proteins. The cDNA sequences are obtained from cDNA libraries, prepared from fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina. The encoded polypeptides have specific uses based on their homology to known polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as lung squamous cell carcinoma of the vulva and gliomas), potent effects on cell growth and development, diseases related to growth or survival of nerve cells including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as a target for anti-tumor drugs. PRO533 may be used in the treatment of Usher Syndrome or Atrophla areata; PRO269 can be used as an anti-thrombotic agent; PRO287 polypeptides and portions may have therapeutic applications in wound healing and tissue repair; PRO317 can be used for treating problems of the kidney, uterus, endometrium, blood vessels, or related tissue, e.g. in the heart of genital tract.

Sequence 552 AA;

Query Match 13.5%; Score 1351.5; DB 20; Length 552;

Best Local Similarity 51.6%; Pred. No. 2.7e-108;  
Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;  
QY 23 STLLTRQAPLPSOKOR-----SFVTRGEPAE- 49  
Db 45 SELLT--AAPLMEQKQWPRALEVDVSVVLLAPPAGMPQSFTHSENDDW 102  
QY 50 GFNHLVVDERTGHIYLGAVNRVYKLSLDKLVLTHTGPDENPKCYPRIVQTNEPLT 109  
Db 103 TENHLTVHQGTGAVVYGAIRVYKLTGNLTIOVAHKTGPEEDNKSRYPLIVQPCSEVLT 162  
QY 110 TNNVNMKLLIYDYNKRLIACSLVQICLKRLLEDLFLKGPYHKEHYLVSNWESGV 169  
Db 163 LTNVNMKLLIYDYNKRLIACSLVQICLKRLLEDLFLKGPYHKEHYLVSNWESGV 222  
QY 170 FGVIVSYNSLNDKLFATADVGKPEYFPTISRRKLTNSEADGMPAYVFHDFVASMKI 229  
Db 223 YGVIVRSEGEDKLFATADVGKPEYFPTISRRKLTNSEADGMPAYVFHDFVASMKI 282  
QY 230 PSDTFIIPDFDIYVYGFSSGNFVYFLTLQPEMVSPPG---STTKQVYVTSKLVRLCKE 286  
Db 283 PSDTLALVSHDFIYVYGFSSGNFVYFLTLQPEMVSPPG---STTKQVYVTSKLVRLCKE 340  
QY 287 DPAFNSYVEVPICGRSGVEYRLLQAAVLSKAGVLRGLGVHPDDLLFTVPSKQGRK 346  
Db 341 DPKFHSYVSLPGCTGAGVYRLLQAAVLSKAGVLRGLGVHPDDLLFTVPSKQGRK 400  
QY 347 MKSLDESALCIFILQINDRIKRLQSCYRGEGTDLMLKVKDIPCSALLTIDNFCG 406  
Db 401 HPPDDLSALCIPFIRAINIQLKRLQSCYRGEGTDLMLKVKDIPCSALLTIDNFCG 460  
QY 407 LDMNAPLGSDVMVRGIPVETEDRDMTSVIAYVYKNSLAFVGTGSKGLKIRVDGPRGN 466  
Db 461 LDINPLGSGTPEVGLTITTSRDMTSVASVYNGYSVVFVGTGSKGLKIRVDGPRGN 514  
QY 467 ALQYE-----TVQVDDGPGVILDMAPSK-DHEOLYIMSERQ 501  
Db 515 ---YEFRCNAIHLLSKESLLEGSYWMFRNYRLQYFLGQR 552

RESULT 15

AAB33420

ID AAB33420 standard; Protein; 552 AA.

XX AC AAB33420;

XX DT 29-JAN-2001 (first entry)

XX DE Human PRO235 protein UNQ209 SEQ ID NO:31.

XX KW Human; immune related disease; diagnosis; antiinflammatory; cardiant;  
KW dermatological; antiarthritic; antirheumatic; immunosuppressive;  
KW haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;  
KW antianemic; hepatotropic; virucide; antipsoriatic; antiallergic;  
KW osteoarthritis; systemic lupus erythematosus; rheumatoid arthritis;  
KW idiopathic inflammatory myopathy; systemic sclerosis; sarcoidosis;  
KW autoimmune vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;  
KW demyelinating disease; immune-mediated renal disease;  
KW inflammatory bowel disease; hepatobiliary disease; Whipple's disease;  
KW autoimmune disease; immune-sensitive enteropathy;  
KW immunological disease; immun-mediated skin disease; allergic disease;  
KW graft rejection; graft-versus-host-disease.

OS Homo sapiens.

XX WO200053758-A2.

XX PD 14-SEP-2000.

XX PF 02-MAR-2000; 2000WO-US05841.

PR 08-MAR-1999; 99WO-US05028.  
PR 10-MAR-1999; 99US-0123618.  
PR 12-MAR-1999; 99US-0123957.  
PR 23-MAR-1999; 99US-0125775.  
PR 23-APR-1999; 99US-0128849.  
PR 20-APR-1999; 99WO-US08615.  
PR 28-APR-1999; 99US-0131445.  
PR 04-MAY-1999; 99US-0132371.  
PR 14-MAY-1999; 99US-0134287.  
PR 02-JUN-1999; 99WO-US12252.  
PR 23-JUN-1999; 99US-0141037.  
PR 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
PR 28-JUL-1999; 99US-0146222.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 05-OCT-1999; 99WO-US23089.  
PR 29-OCT-1999; 99US-0162506.  
PR 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
PR 30-NOV-1999; 99WO-US28409.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28564.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US31274.  
PR 30-DEC-1999; 99WO-US00219.  
PR 05-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.

(GETH ) GENENTECH INC.

Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;  
Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;  
Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;

WPI; 2000-572271/53.

N-PSDB; AAC58585.

Sixty four PRO polypeptides, useful in the diagnosis and treatment of  
immune related disorders, e.g. systemic lupus erythematosus, rheumatoid  
arthritis, osteoarthritis, thyroiditis and diabetes mellitus -

Claim 33; Fig 14; 309pp; English.

The present invention describes sixty four human PRO proteins which can  
be used in the treatment of immune related diseases. The human PRO  
proteins, anti-PRO antibodies, agonists and antagonists are useful for  
treating and diagnosing immune related disorders. The disorders are  
selected from systemic lupus erythematosus, rheumatoid arthritis,  
osteoarthritis, juvenile chronic arthritis, spondyloarthritis,  
systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's  
syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic  
anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,  
immune-mediated renal disease, demyelinating diseases of the central  
and peripheral nervous systems, hepatobiliary diseases, inflammatory  
bowel disease, gluten-sensitive enteropathy and Whipple's disease,  
autoimmune or immune-mediated skin diseases, allergic diseases,  
immunological diseases of the lung, and transplantation associated  
diseases including graft rejection and graft-versus-host-disease.  
AAC58397 to AAC58578 represent PCR primers and hybridisation probes used  
in the isolation of human PRO sequences. AAC58579 to AAC58642 and  
AAB33414 to AAB33477 represent human PRO polynucleotide and protein





Query Match 13.58; Score 1349; DB 4; Length 1568;  
Best Local Similarity 24.88; Pred. No. 5.8e-112; Indels 576; Gaps 74;  
Matches 480; Conservative 291; Mismatches 591;

QY 72 YKLSDDLKVLVTHETGDEDPNPKYPRIVQTCNEPLTTT-----NNVNKMLLDYK 123  
DB 70 YSLEHSLRLYRDOAG-----NCTEPVSLAPPARPRGSSFSK-LLLPYR 113  
QY 124 ENR-----LIAGCSLYQGI CKLLRLBDFLKLGPYHKKHYLSGV-----NESGSPGV 172  
DB 114 EGAAGLGLLLTGWTDFRGACEVRPLGNL-----SRNSLRNGTEVWSCHPOGSTAGV 165  
QY 173 IVSYSNLDDKLFIA-----TAVDGKPEVFPPTISSRKLTKNS 208  
DB 166 VYR-AGRNWRYLAAVATYVLPPEBTASRCPAASDHDHTALDKDTRSLATQELGRLK 224  
QY 209 EADGMFAYVFHDFV-ASMIKIPSDTFTIIPDFDIYVYGFSSGNFVYFILTQPEMVSP 267  
DB 225 LCEGAGSLHVDFAFLMNGSIYFP-----YYPNYVTSG-----AATGWSPMARIA 268  
QY 268 GSTTKEQVYTKLVRCKEDTAFNSYVEVIGCBRSQVE-YRLLQAAVLSKAGAVLGRTL 326  
DB 269 QST-----EVLFGQ--QASLDGCHGHPDGRLLLSLVEA----- 302  
QY 327 GVHPDDDLLETFVS--KGOKKMSLDESALCIFILKQINDRIKERLQSCYRGEGTDLA 384  
DB 303 -----LDWAGVSAAGGQERSPTTALCLFPMSEIOARAK-----VS 344  
QY 385 W-LKVKDIPCSSALLTIDDFNFCGLDMNAPLGVSDMRGIPVFTEDRMTSVIAVYKNH 443  
DB 345 WDFKTAESHCKEG-----DQP-----ERVQPIASSTLIHSDLTSVYGVVWNR 387  
QY 444 SLAFVGTSGKLKIRVDGPRGNALQVETQVV-----DPGVLRDWARSKD 490  
DB 388 TVLFLGTGDOQLLVIL-----GENLTSNCPVEIYEKEETPVFKLVDPDPV-----K 435  
QY 491 HEOLYIMSERQLTRVPVESCQYQSCGECGLSGDPHCGWCLVHNTCTRKERCERSKEPRR 550  
DB 436 NIYIYLTAGREVRIRVANCNKHKSCECLTATDPHCGWCHSLORCTFOGDC-----487  
QY 551 FASEMKOCVRLTHVPNNISVSQYNVLLVLEFYVNPVPELSAGVNTCFEDLSEMDGLVGNQI 610  
DB 488 -----VHSENLE-----NWLIDISSG----- 502  
QY 611 QCYSPAKEYPRIITENGDDHVVQLQKSKETGTMFTASTSPFYNCVHNSCL-SCVESP 669  
DB 503 -----AKCPKI-----QIIRSEKRTVTVMVGSF-----SPRHSKCMVKNVDSS 542  
QY 670 YRCHWCKYRHVCTHDPKTCSPQEGRVKLPDPCQLLRVDKILVPVEVIKPTLKAQNLPO 729  
DB 543 -----REL----- 551  
QY 730 PQSGQRGYECILNTQSGEQRPALRFNSSVQONTYSYEGMEINNLPLVELTVVWN--G 787  
DB 552 PNR-----TCTCSI-----PTRAITYKDVSVVNVWFSFGS-----WNLS 585  
QY 788 HFNIDNPAQNVHLYKCGAMRESGLCLKADPDPAFCWCGQPGQCTLRQHCPAESOWLE 847  
DB 586 RPNFTN-----CSSLKE-CPACVET-----GCWCKSARRC----- 615  
QY 848 LSGAKSKTNPRITEIIPVTGPRGGTKVTIRGENLGLFPRDIAASHVKVAGVGCSPVDG 907  
DB 616 -----IHPFTA----- 625  
QY 908 YIPAEQIVCEM-----GEAKPSOHAGVEICVAVCRPEFMARSSQLYVFTLTLSDLP 961  
DB 626 DYERNQPCPVAVEKTSGGGRPKENKG-----NRTNQLQVFI-----IKSIEP 669  
QY 962 SRGPMSGTQVITIGTNLNAQSNV-VVMFGKQPC-----LFHRRSPSYIVCNVTTSSDEV 1014  
DB 670 QKSTLGSNVIVTGANFTRASNITMILKGTSTCDKQVIOVSHVNDTHMKFSLPSSRK- 728

QY 1015 LEMK-VSVQVDRAKIHQDLVFQYVEDEFTVIRIEPWSIVSGNTPIAVWGTHTLDLQNPQI 1073  
DB 729 -EMKDVCIQPDGNGCSSVSGLSYIALPHCSLIIPATTWISGQONITMGRNFDVIDN--L 785  
QY 1074 RAKHGGEKHINI CEVLNATEMTCOAPALAGPDHQSOLDTERPEEPFGFILDNVQSLLILNK 1133  
DB 786 IISHELKGNINVSVCVATYCGFLAPSL-----KSSKVRTNVTVKLRVD 830  
QY 1134 T-----NFTYYPNVEFAFGPSGILELKPGPPIILKGNLIPPVAGGVNKNLYTVLVGSK 1188  
DB 831 TYLDCGTLQVREDPRFTGYR-----VESEVTELEVKIO-----KENDNFNISK 875  
QY 1189 PCTVTV---SDVQLLCESPNI-----GRHKVMARVGMGEYS 1222  
DB 876 DIEITLPHGNGQLNCSPENITRNQDLFTILCKIKIGIKTASTIANSSKKVRKLGLE-- 933  
QY 1223 PGWYIAPDS-PLSLPAIVSITAVAGGLLIIFIVAVLIAYKRSRESDLTKELQOMQMDNL 1281  
DB 934 ---LYVEQESVPSTWYFLIVLPV---LLVIVIFAAGVTRHKSKE--LSRKQSQ-OLELL 984  
QY 1282 ESRVALECKEAFABELOTDIHELTSDLDCAG-IPFLDYRTYTMRVLPF--GIEDHPVLRLD 1338  
DB 985 ESELRKEIRDFABELQMDKLDV---VDSFGTVPFLDYKHPALRTEFPESGGTHIFTEDM 1041  
QY 1339 EPGYQROERVEKGLKIFA--OLINNKVFLSFIRLESQSFMSMDRGNVASLIMTVLOS 1396  
DB 1042 ---HNRDANDKNESLTALDALICNKSFLVTVIHTLEKQKNFVKDRCLIFASFLATLQ 1097  
QY 1397 KLEYATDVLKOLLADLIDKNLESKNHPKLLARRRESVAEKMILTNMFTFLLYKFLKECAGE 1456  
DB 1098 KLVYLTILEVLTDLMEQC--SNMQPKMLRRTESVVEKLLTNMWSVCLSGFLRETVE 1155  
QY 1457 PLFLSFCALQOMKEGPIDAITGBARYSLSBDKILRQOQIDYKTVLSCV---SPONANSP 1513  
DB 1156 PFYLLVTLTNGKINKGPDVITCKALYTLNEDWLLWQVPESTVALNVVFEKIPENESAD 1215  
QY 1514 ---EVPVKILNCDTITQVKEKILDAIFKNVPCSHRPAKADMDLEWRQSGARMILQDEDI 1570  
DB 1216 VCRNISVNLVDCDITGOAKEKIFQAFLSKNGSPYGLQLEIGLELQMGTRQKELLIDSS 1275  
QY 1571 TTKLENDKRLANTLAHVQPDGVSVALVSKQVTA VNAVNNSTVSRTSASKYENMIRYTCG 1630  
DB 1276 SVILEDGITKNTLGHVEISNGSTIKV-----FKIANFTSD 1312  
QY 1631 PDSLRSPMTITPDLESQVGMHLVKNHEHGDQKEGRGSK--MVSEIYTLRLATKGT 1688  
DB 1313 VEYSDDHCHLILPSEA---FQDVQGRH-----RGKHFKVKEMWYTLKLLSTKVAI 1361  
QY 1689 QKFVDDLPETIFSTAHRGSAIPALIKYMPDFDLDEQADKHGHDHVRHTWKNCLPLRFW 1748  
DB 1362 HSVLEKLFRSIWSLPN--SRAPPAIKYFPDFDLDAQAKNKITDPPDVVHTWKNLSPLRFW 1419  
QY 1749 VNMILKNPQVFDIHKNSITDACLVSVAOTFMDSCSTSEHRLGKDSPSNKLAKDIPSYK 1808  
DB 1420 VNILKNPQVFDIHKNSITDACLVSVAOTFMDSCSTSEHRLGKDSPSNKLAKDIPSYK 1808  
QY 1809 NWVRYRYSDIGKMPAISQDMNAYLAEQSRMHMNEFNTMSALSSEIFSYVGYKSEELGPL 1868  
DB 1480 EEVKSYYKAIKRDPLPLSSSEMEEFITQESKXENEFNEBEVALTEIVKYIVKYFDEILNKL 1539  
QY 1869 DHD---DOCGKQKAYKL 1883  
DB 1540 ERERGLEEAQQLLHVKV 1557

RESULT 2  
US-09-458-791-2  
; Sequence 2, Application US/09458791  
; Patent No. 6174689  
; GENERAL INFORMATION:  
; APPLICANT: Spriggs, Melanie  
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
; RECEPTOR DNA AND POLYPEPTIDES

NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Janis C. Henry  
STREET: 51 University St.  
CITY: Seattle  
STATE: WA  
COUNTRY: US  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS/Windows 95  
SOFTWARE: Word for Windows 95, 7.0a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/458,791  
FILING DATE: 10-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/958,598  
FILING DATE: 28-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Henry, Janis C  
REGISTRATION NUMBER: 34,347  
REFERENCE/DOCKET NUMBER: 2631  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)470-4189  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1569 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-458-791-2

Query Match 13.5%; Score 1349; DB 4; Length 1568;  
Best Local Similarity 24.8%; Pred. No. 5.8e-112;  
Matches 480; Conservative 291; Mismatches 591; Indels 576; Gaps 74;

436 NIYVLTAGKEVRRIRVANCNKHSCECLTATDPHCGWCHSLQRCCTFGQDC----- 487  
551 FASEMKQCVRLTVHPNNISVSQYNVLLVLETVNVPBLSAGVNCCTFEDLSEMDGLVGNQI 610  
488 -----VHSENLE-----NWLDISSG----- 502  
611 QCYSPAAKEVPRIITENGDDHHVVQLQKSKETGTMFASTSFVYNCVHNSCL-SCVESP 669  
503 -----AKCPKI-----QIIRSSKEKTVTMVGSF-----SPRHSKCMKVNVDSS 542  
670 YRCHWKYRHVCHDPTKCSFQEGRVKLPEDCPQLLRVDKILVPVEVIKPTLKAQLPQ 729  
543 -----RELQ----- 551  
730 PQSGQRGYECILNIQSGEQRPALRFNSSSVQCONTSYSEYEGMEINNLVELTVVMN--G 787  
552 PNR-----TCTCSI-----PTRATYKDVSVVNVMFSGS-----WNLSD 585  
788 HFNIDNPAQNVKHLKYGAMRESGLCLKADPDPAFCGCGQCGQCTLRQHCPAESQWLE 847  
586 RNFNTN-----CSSLKE-CPACVET-----GCWACKSARRC----- 615  
848 LSGAKSKTNPRITEIIPVTGPREGGKVTIRGENLGLEFRDIASHVKVAGVGCSPLDVG 907  
616 -----IHPFTA----- 625  
908 YIPAEQIVCEM-----GEAKPSQAHGVEICVAVCRPEFMARSSQLYFMTLTSLDKP 961  
626 DYERNEQCPVAVEKTSGGGRPKENKG-----NRTNQLQVYF-----IKSTEP 669  
962 SRGPMGGTQVITGTNLAGSNV-VVMFGKQPC-----LFHRRSPSYICNTTSSDEV 1014  
670 QKYSTLGKSNVITGANFTASNIWILKGTSTCDKDVIOVSHVNLNTHMKFSLPSSRK- 728  
1015 LEMK-VSVQVDRAKIHQDLVFQVEDPTVIRIPEWSIVSGNTPIAVGTHLDLIONQOI 1073  
729 -EMKDVCIQPDGNCSSVSGSLSYIALPHCSLIIPATTWISGQGNITMMGRNFDVIDN--L 785  
1074 RAGHKGHEINICEVLNATEMTQAPALALGPDHQSOLDTERPEFGFILDNVQSILLANK 1133  
786 IISHELKGNINSEYCVATYCGFLAPSL-----KSSKVRTNVTVKLRVQD 830  
1134 T-----NFTVYPPVFEAFPGSGILELKPCTPIILKGNLIPPVAGGNVKNLYTVLGEK 1188  
831 TYLDCGTLQVREDPRFTGYR---VSEVDTELEVKIQ-----KENDNFNISK 875  
1189 PCTVTV---SDVQLCESPNLI-----GRHKVMARVGGMEYS 1222  
876 DIEITLPHGNGQLNCSFENITRNQDLTTILCKIKIKTASTIANSSKKVRVKGLENLE-- 933  
1223 PGMVYIAPDS-PLSLPAIVSIAVAGGLIIFIVAVLIAYKESRESDLTLKRLQWMDNL 1281  
934 ---LYEQESVPTWTVFLVLPV---LLVIVIFAAGVTRHKSKE--LSRKQSQ-QLELL 984  
1282 ESRVALECKEAFABLOTDIHELTSIDLPGAG-IPPLDYRTVTRVLPF--GIEDHPVLRDL 1338  
985 ESELKEIRDGPAELQMDKLDV---VDSTGTVPFLDYKHFAIRTFPFESGGFTIIFTEDM 1041  
1339 EPGYQERVEKGLKFA--OLINNKVFLSFIIRLTESQRSFSMRDRGNVASLIMTVLQS 1396  
1042 ---HNRDANDKNESLTALDALLICNSFLVTIHTLEKQKNSVKDRCLFASFLTIALQT 1097  
1397 KLEYATDVLKOLLADLIDKNLESKNHKLIRRTESVAEKMWTNFWFLLYKFLKECAGE 1456  
1098 KLVYLTSLILEVLRDLMEQC--SNMQPKMLRRTESVVEKLLTNWMSVCSGFLRETVE 1155  
1457 PLFSLFCAIKQOMEKGPIDAITGEARYSLSEDKLIRQOIIDYKTLVLSCV---SPDNANSP 1513  
1156 PFYLLVTTLNQKNGPVDVITCKALYTLNEDWLLWQVPFESTVALNVVPEKIPENESAD 1215  
1514 ---EVPVKILNCPTITQVKEKILDAIFKNVPCSHRPKAADMDLEWROGSGARMILQDEDI 1570

Query Match	13.5%;	Score 1349;	DB 4;	Length 1568;
Best Local Similarity	24.8%;	Pred. NO. 5.8e-112;		
Waste	400;	Conservative	291.	Mismatches 591;
Indels	576;			Indels 576;
Caps	74;			Caps 74;





QY	692 E 692	
Db	677 K 677	
RESULT 7		
US-08-833-391-62		
/ Sequence 62, Application US/08833391		
/ Patent No. 6013781		
/ GENERAL INFORMATION:		
/ APPLICANT: Goodman, Corey S.		
/ APPLICANT: Kolodkin, Alex L.		
/ APPLICANT: Matthes, David R.		
/ APPLICANT: Bentley, David R.		
/ APPLICANT: O'Connor, Timothy		
/ TITLE OF INVENTION: The Semaphorin Gene Family		
/ NUMBER OF SEQUENCES: 100		
/ CORRESPONDENCE ADDRESS:		
/ ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP		
/ STREET: 268 Bush Street, Suite 3200		
/ CITY: San Francisco		
/ STATE: CA		
/ COUNTRY: USA		
/ ZIP: 94104		
/ COMPUTER READABLE FORM:		
/ MEDIUM TYPE: Floppy disk		
/ COMPUTER: IBM PC compatible		
/ OPERATING SYSTEM: PC-DOS/MS-DOS		
/ SOFTWARE: Patent In Release #1.0, Version #1.25		
/ CURRENT APPLICATION DATA:		
/ APPLICATION NUMBER: US/08/833,391		
/ FILING DATE:		
/ CLASSIFICATION: 530		
/ PRIOR APPLICATION DATA:		
/ APPLICATION NUMBER: US 08/121,713		
/ FILING DATE: 13-SEP-1993		
/ ATTORNEY/AGENT INFORMATION:		
/ NAME: Osman, Richard A.		
/ REGISTRATION NUMBER: 36,627		
/ REFERENCE/DOCKET NUMBER: B94-002-1		
/ TELECOMMUNICATION INFORMATION:		
/ TELEPHONE: (415)343-4341		
/ TELEFAX: (415) 343-4342		
/ TELEX:		
/ INFORMATION FOR SEQ ID NO: 62:		
/ SEQUENCE CHARACTERISTICS:		
/ LENGTH: 724 amino acids		
/ TYPE: amino acid		
/ TOPOLOGY: linear		
/ MOLECULE TYPE: protein		
US-08-833-391-62		
Query Match 2.3%; Score 227; DB 3; L		
Best Local Similarity 20.4%; Pred. No. 4.4e-11;		
Matches 147; Conservative 118; Mismatches 260;		
QY	53	HLVVDERTGHYLGAVNRYIKLSSDLKVLVTHETGPDENN
Db	72	HNNEDRDT-LYVGAMDRVFRV--NLQNISSNCNRDAIN
QY	113	NVNMKLLIDYKENRLIACSLYQGICKLRLRLDLFKLGEPI
Db	120	SKGKSQIFDCK-NHVRVITQSDMQG-----DRLVYVCGTN
QY	168	SVFGVIVS-----YGNLDDKLFIAIVATV---DGKPEYFFPTI
Db	172	YVIGVGLGIKACPYDPLDN---STAIVYVENGNGPGLPGL
QY	220	DFEVASMIKIPSDTFTIIPDPDI-----YYVYGFSSGNF
Db	225	TDLVNTSAKRLEYKFKRTLKYDSKRLDKPKNFVGSFDIGEY

Db 280 GKAVY-SRIARVCKDVGKGNLHAWNATYKARLNCISGEFFPYENEIOQVY 332  
QY 321 VLGRITGVHPDDLLFTVFSKQKRMKSLDEGALCIFILKQI-----NDRIKERLQCYR 376  
Db 333 -----QLPSDKSRFFATFT-----TSTNGLIGSAVCSFHINEIQAAFNGKFKQSSS--- 379  
QY 377 GEGTLDLAWLKVD--IPCSSALLTIDNFCGLDMNAPLGVSDMVRGIP----- 423  
Db 380 -----NSAWLPVLNSRVPEPRPGTCVNDT-----SNLPDVTVLNFIKSHPLMDKAVNHEHN 429  
QY 424 -----VFTE-----DRDMTSVIAYVYKNSHSLAFVGTGSKGLKKIRVDGPRGNAL-QY 470  
Db 430 NPVYKRDVLVTKLVVDKIR-----IDILNQEYIVYVGTNLGRIYKIVQYRNGESLSKL 485  
QY 471 ETQVQVDFPVLDMAFSKDHEQLYIMSERQITRVPVESCG-QYQSCGCEGLSGDPCGHW 529  
Db 486 LDIFEVAPNEAIQVMEISQTRKSLYIGTDHRIKQIDLAMCNRRYDNCFRCV--RDPYCGW 543  
QY 530 CVLHNTCTRKER-----CERSKEPRRPFASEMKQCVRLTVHPNPNISVSQYNVL 576  
Db 544 DKEANTCRPYELDLLOQVANETSICDSSVLKKIVVYQSVHLGCF----- 591  
QY 577 LVLETYNVPELSAGVNCFTFEDLSMDGLVGNQIOCYSPAAREVPRIITENGHHVVLQ 636  
Db 592 -----VKIPEVLKNEQVTWYHHSKDKGRY-----EIR-YSPTK-----YIETTERGLVWVS 636  
QY 637 LKSKETGTMFTASTSFVYNCVSHNSCLSCVSPYRCHWKYR-----HVCTHDPKTCFQ 691  
Db 637 VNEADGGR-----YDCHLGGSL-----CSYNTITVDAHRCTPPNKSNDYQ 676  
QY 692 E 692  
Db 677 K 677

RESULT 8

US-09-060-610-62  
; Sequence 62, Application US/09060610  
; Patent No. 634544  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S.  
; APPLICANT: Kolodkin, Alex L.  
; APPLICANT: Matthes, David  
; APPLICANT: Bentley, David R.  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/060,610  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/835,268  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B94-002-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)343-4341

TELEFAX: (415) 343-4342  
TELEX:  
; INFORMATION FOR SEQ ID NO: 62:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 724 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-060-610-62  
Query Match 2.3%; Score 227; DB 4; Length 724;  
Best Local Similarity 20.4%; Pred. No. 4.4e-11;  
Matches 147; Conservative 118; Mismatches 260; Indels 196; Gaps 36;  
QY 53 HLVDERTGHIYLGAVNRIYKLSLDKLVLVTHETGPDENPKCYPPRIVQTCNEPLTTN 112  
Db 72 HNNEDRT--LYVGAMDRVPRV--NLQNISSNCNRDAIN-----LEPTRDDVWSCV 119  
QY 113 NVNQMLLDYKENRLIACGSLYQIGICKLLRLLEDLKLGEYPVHKKEHY-----LSGVNESH 167  
Db 120 SKGSKQIFDCK-NHVRVIQSDQG-----DRLYVCGTNAHNPKDYIYIANLTHLPRSE 171  
QY 168 SVFGVIVS-----YSNLDLDFIATAV---DGKPEYFPTISSRKLTKNSEADGMFAYVFH 219  
Db 172 YVIGVGLGIAKCPDPLDN---STAIYVNGNPGGLPGLYSTGTTNAEFTKAD---TVIFR 224  
QY 220 DEFVAMIKIPSDFTTIIIPDFI-----YVYGFSSGNFVYFLTLP--EMVSPPGSTT 271  
Db 225 TDLYNTSAKRLYKPKRTLKYDSKWLDPKPNFSGFDIGYVYFFPRETAVEXIN-----C 279  
QY 272 KEQVYTSKLVLCXEDTA-----FNSYVBPVIGCERSG---VEYRLLOAAVLSKAGA 320  
Db 280 GKAVY-SRIARVCKDVGKGNLHAWNATYKARLNCISGEFFPYENEIOQVY 332  
QY 321 VLGRITGVHPDDLLFTVFSKQKRMKSLDEGALCIFILKQI-----NDRIKERLQCYR 376  
Db 333 -----QLPSDKSRFFATFT-----TSTNGLIGSAVCSFHINEIQAAFNGKFKQSSS--- 379  
QY 377 GEGTLDLAWLKVD--IPCSSALLTIDNFCGLDMNAPLGVSDMVRGIP----- 423  
Db 380 -----NSAWLPVLNSRVPEPRPGTCVNDT-----SNLPDVTVLNFIKSHPLMDKAVNHEHN 429  
QY 424 -----VFTE-----DRDMTSVIAYVYKNSHSLAFVGTGSKGLKKIRVDGPRGNAL-QY 470  
Db 430 NPVYKRDVLVTKLVVDKIR-----IDILNQEYIVYVGTNLGRIYKIVQYRNGESLSKL 485  
QY 471 ETQVQVDFPVLDMAFSKDHEQLYIMSERQITRVPVESCG-QYQSCGCEGLSGDPCGHW 529  
Db 486 LDIFEVAPNEAIQVMEISQTRKSLYIGTDHRIKQIDLAMCNRRYDNCFRCV--RDPYCGW 543  
QY 530 CVLHNTCTRKER-----CERSKEPRRPFASEMKQCVRLTVHPNPNISVSQYNVL 576  
Db 544 DKEANTCRPYELDLLOQVANETSICDSSVLKKIVVYQSVHLGCF----- 591  
QY 577 LVLETYNVPELSAGVNCFTFEDLSMDGLVGNQIOCYSPAAREVPRIITENGHHVVLQ 636  
Db 592 -----VKIPEVLKNEQVTWYHHSKDKGRY-----EIR-YSPTK-----YIETTERGLVWVS 636  
QY 637 LKSKETGTMFTASTSFVYNCVSHNSCLSCVSPYRCHWKYR-----HVCTHDPKTCFQ 691  
Db 637 VNEADGGR-----YDCHLGGSL-----CSYNTITVDAHRCTPPNKSNDYQ 676  
QY 692 E 692  
Db 677 K 677

RESULT 9

PCT-US94-10151A-62  
; Sequence 62, Application PC/TUS9410151A  
; GENERAL INFORMATION:  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSER: FLEHR HOBBACH TEST ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/10151A  
FILING DATE: 13-SEP-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299 FHT UR  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 724 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-10151A-62

Query Match 2.3%; Score 227; DB 5; Length 724;  
Best Local Similarity 20.4%; Pred. No. 4.4e-11; Indels 196; Gaps 36;  
Matches 147; Conservative 118; Mismatches 260;  
53 HLAVDERTGHIYLGAVNRITKLSLKLVLVTHETGDEDNPKYPPRIVQTGNEPLTTTN 112  
72 HNVEDRT--LVGAMDRFRV--NLQNSSNCNRDAIN-----LEPTRDDVVSCV 119  
113 NVNKMILLIDYKENRLIACGSLVQGIKLLRLDLFKLGEPPYKHKEHY-----LSGVNESH 167  
120 SKGKQIFDCK-NHVRVIOQMDQ-----DRLYVCGTNAHNPDKYVIYANLTHLPRSE 171  
168 SVFGVIVS-----YSNLDDKLFATAV---DGKPEYFPTISSRKLTKNSEADGMFAVVEH 219  
172 YVIGVGLGIKACPYDPLDN-----STAIYVNGNPGGLGLYSGTNAEFTKAD---TVIFR 224  
220 DEFVASMIPKIPSDTFTIIPDFDI-----YVYVGFSSGNFVYPLTLQP--EMVSPGSGTT 271  
225 TDLYNTSAKRLVKFKRTLKYDSKWLDPKFNFGVDFIGEVYFFPRETAVEYIN-----C 279  
272 KEQVTSKLVRLCKETA-----FNSVVEVPIGCSRG---VEYRLLOAYLSKAGA 320  
280 GRAVY-SRIARVCKDVGKGLLAHNAWATYLRKLNCSISGEPPFYFNEIQSVY----- 332  
321 VLGRITGVHPDLDLFTVFSKQKRWKSLDESALCIFILKQI-----NDRIKERLOSRYR 376  
333 -----QLPSSDKSRFPATFT-----TSTNGLISAVCSFHINEIQAAFNGKFEQSSS--- 379  
377 GEGTLDLAWLKVYKD--IPCSSALLTIDDDNFCGLDMNAPLGVSMDVRGIP----- 423  
380 -----NSAWLPVLSNRVPRPCTGVNDT-----SNLPDTLVNFIHPLMDKXVNHEN 429  
424 -----VFTE---DRDMTSVIAVYKYNHSLAFVGTGSKGLKIRVDGPRGNAL-QY 470  
430 NPVYVKRDLVFTKLVDKIR-----IDILNQEYIVYVGTGNLGRYIKIVQYVRNGESLSKL 485  
471 ETVOVVDPGVLRDMAFSKDHEOLYIMSERQLTRVPVESCQ-QYQSCGCLSGSDPHCGW 529  
486 LDIFEVAPNEAIQWMEISQTRKSLYIGTDHRIKQIDLAMCNRRYDNCRCV--RDPYCGW 543

530 CVLHNTCTRKER-----CERSKPRPRFASFMKQCVRLLTVHPNNISVSQYNVL 576  
544 DKEANTCRPYELDLQDVANETSICDSSVLKKKIVVTYQSVHLGCF----- 591  
577 LVLETYNVPELSAGVNCITPEDLSMDGLVGVNQICQYSPAACEVPRITENGDDHHVQLQ 636  
592 -----VKIPEVLKNEQVTVYHHSKDKGRY---EIR-YSPTK-----YIETTERGLVVVS 636  
637 LKSKETGMTFASTSFVFPYNGSVHNSCLSCVSPYRCHWCKYR-----HVCTHDPKTCFQ 691  
637 VNEADGR-----YDCHLGSSLL-----CSNITVDAHRCTPPNKSNDYQ 676  
692 E 692  
677 K 677  
RESULT 10  
US-08-121-713D-58  
; Sequence 58, Application US/08121713D  
; Patent No. 5639856  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S.  
; APPLICANT: Kolodkin, Alex L.  
; APPLICANT: Matthes, David  
; APPLICANT: Bentley, David R.  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/121,713D  
; FILING DATE: 13-SEP-1993  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B94-002-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 730 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-121-713D-58

Query Match 2.1%; Score 207; DB 1; Length 730;  
Best Local Similarity 18.9%; Pred. No. 2.9e-09;  
Matches 111; Conservative 107; Mismatches 219; Indels 150; Gaps 27;  
43 ERGEPAEGFNHLVVDERTGHIYLGAVNRITKLS-SDL-----KVLVTHETGDEDNPKY 96  
41 FLGNESHKDHFKLLEKDHNSLLVGCARNIVYNISLRDLTFTEQRIEWHSSGAHRE--LCY 98  
97 -PRIVQTCNEPLTTTNVNNKMLLDYKENRLIACG-SLYQGIKLLRLDLFKLGEYPH 154  
99 LKKGSEDDCQ-----NYIRVLAKID--DDRVLICGTNAYKPLCRHVALKD----GDYVV 146

RESULT 11  
US-08-835-268-58.  
; Sequence 58, Application US/08835268  
; Patent No. 5807826  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S.  
; APPLICANT: Kolodkin, Alex L.  
; APPLICANT: Matthes, David  
; APPLICANT: Bentley, David R.  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/835,268  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/121,713  
; FILING DATE: '13-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B94-002-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)343-4341  
; TELEFAX: (415) 343-4342  
; TELEFAX:  
; INFORMATION FOR SEQ ID NO: 58:

```

RESULT 12
US-09-060-692-58
; Sequence 58, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

RESULT 13  
US-08-833

QY 374 CYRGETLIDLAWLKVKDI-----PCSSALLTIDNFCGLDMNAPLGVSMDVRGIPVF 425  
Db 331 PFKEQETMNSNWLAVPSLKVPKPGQCVNDSRTLPD-----VSVNFVKSHTLMDAEPAPF 386  
QY 426 TEDDRMTSVIAVY-----KNHSLAFVGTGSKGLKKI-----RVD 461  
Db 387 FTRPILIRISLOVREFTKIAVDQOVRTDGRKAYDVLFGTDDGKVIKALNSASPDSSDSTD 446  
QY 462 GPRGNALQVETQVVDGPGVLRDM---AFSKDHEQLYIMSERQLTRVPVSCG--QYQSC 516  
Db 447 -----SVVIEELQVLPVGVFNKLVYVRMDGDDSKLVVSDDEILAIKLRHRCGSDKITNC 501  
QY 517 GECIGSDPHCGWCVLHNTCTRKERCERSKEPRF-----ASEMKOC 558  
Db 502 RECVSLODPYCAWMDNVELKCTAVGSPDWSAGKRFRFQNIISLGEHKAC 548

RESULT 14  
US-09-060-610-58  
; Sequence 58, Application US/09060610  
; Patent No. 6344544  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S.  
; APPLICANT: Kolodkin, Alex L.  
; APPLICANT: Matthes, David R.  
; APPLICANT: Bentley, David R.  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/060,610  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/835,268  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B94-002-1  
; TELEPHONE: (415)343-4341  
; TELEFAX: (415) 343-4342  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 730 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-060-610-58

Query Match 2.1%; Score 207; DB 4; Length 730;  
Best Local Similarity 18.9%; Pred. No. 2.9e-09;  
Matches 111; Conservative 107; Mismatches 219; Indels 150; Gaps 27;  
QY 43 FRCEPAEGFNHLYVDERTGHIYLGAVNNRIYKLS--SDL-----KVLVTHETGDPEDNPKCY 96  
Db 41 FLGNESHKDFKLEKDHNSLLVGARNIVNLSRLDTEFTQRIEWHSSGAHRE--LCY 98  
QY 97 -PPRIVOTCNEPLTTTNNVKNMLLDYKENRLIACG-SLYQGICKLRLLEDLFLKGEPPYH 154

Db 99 LKGSSEDDCQ-----NYIRVLAKID--DORVLI CGTNAYKPLCRHYALKD-----GDYVV 146  
QY 155 KKE-----HYLSGVNWSGVFGVIVS--YSLNDDDKLFATATAVDGKPEYEPPTISS 201  
Db 147 KEYEGRGLCPFPDPDHNSTAIYSEGQLYSATVAFSGTIDPLIYRG-----PLRTE 196  
QY 202 RKLTKNEADCMFAYVPHDFVASMIKIPSDTFTIIPDFIYVYVFGSGNPFVFLTLQP 261  
Db 197 RSDLKQLNAPNFVNTMEYNDFI-----FFFFRETAVEYI---NCGRKIY----- 237  
QY 262 EMWSPPGSTTKQVQYVTKLVRCLKEDTA-----FNSYVEVPICGERSGVEYRLLQAA 313  
Db 238 -----SRVARVCKHDKGPHQGGDRWTSFLKRLNCSVPG--DYFFYFNE 280  
QY 314 YLSKAGAVLGRTLGVHPDDDLFTVFSKGQRKMKSLDESALCIFIILKQINDRIKERLOS 373  
Db 281 IQSTSDIIEGNYGG--QVEKLIYGVFT---TPVNSIGGSVAVCAFMSKS-----ILESFDG 330  
QY 374 CYRGETLIDLAWLKVKOI-----PCSSALLTIDNFCGLDMNAPLGVSMDVRGIPVF 425  
Db 331 PFKEQETMNSNWLAVPSLKVPKPGQCVNDSRTLPD-----VSVNFVKSHTLMDAEPAPF 386  
QY 426 TEDDRMTSVIAVY-----KNHSLAFVGTGSKGLKKI-----RVD 461  
Db 387 FTRPILIRISLOVREFTKIAVDQOVRTDGRKAYDVLFGTDDGKVIKALNSASPDSSDSTD 446  
QY 462 GPRGNALQVETQVVDGPGVLRDM---AFSKDHEQLYIMSERQLTRVPVSCG--QYQSC 516  
Db 447 -----SVVIEELQVLPVGVFNKLVYVRMDGDDSKLVVSDDEILAIKLRHRCGSDKITNC 501  
QY 517 GECIGSDPHCGWCVLHNTCTRKERCERSKEPRF-----ASEMKOC 558  
Db 502 RECVSLODPYCAWMDNVELKCTAVGSPDWSAGKRFRFQNIISLGEHKAC 548

RESULT 15  
PCT-US94-10151A-58  
; Sequence 58, Application PC/TUS9410151A  
; GENERAL INFORMATION:  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR HOHBACH TEST ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/10151A  
; FILING DATE: 13-SEP-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299 FHT UR  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 730 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

PCT-US94-10151A-58

```
Query Match          2.1%; Score 207; DB 5; Length 730;
Best Local Similarity 18.9%; Pred. No. 2.9e-09;
Matches 111; Conservative 107; Mismatches 219; Indels 150; Gaps 27;

QY 43 FRGEPAAGNHLVDEBTHIYLGAVNRIYKLS-SDL-----KVLVTHETGPDEDNPKCY 96
Db 41 FLGNESHKDHFKLEKDHNSLLVGNRIYVNIISRLDTEFTEQRIEWHSSGAHRE--LCY 98
QY 97 -PPRIVQTCNEPLTTNNVKNMLLDYKENRLIACG-SLYQGICKLLRLLEDLFKLGEPYH 154
Db 99 LKGSSEDDCQ-----NYIRVLAKID--DDRVLICGTNAYKPLCRHYALKD-----GDYV 146
QY 155 KKE-----HYLSGVNESGSVFGVIVS-VSNLDDKLFATADVCKPEYFTISS 201
Db 147 EKEYEGRGLCPDPDHNSTAIYSEGOLYSATVADFSGTDPLIYRG-----PLRTE 196
QY 202 RKLTKNSEADGMFAVYFHDHFVASMIKIPSDFTTIIIPFDIYVYVGFSSGNFVYFLTLP 261
Db 197 RSDLKQLNAPNEVNTMEYNDFI-----FFFRETAWEYI--NCGKAIY----- 237
QY 262 EMVSPPGSTTKQOVVTSKLVRLCKEDTA-----FNSYVEVPICGERSGVEYRLLOAA 313
Db 238 -----SRVARVCKHDKGPHQGGRWTSFLXSRNLNCVPG-DYPFYFNE 280
QY 314 YLSKAGAVLGRTLGVHPDDLLFTVFSKGQKRMKSLDESALCIFILKQINDRIKERLOS 373
Db 281 IQSTSDIIEGNYGG--QVEKLIYGVFT-----TPVNSIGGSVAVCAFMSKS----ILESFDG 330
QY 374 CYRGEGTLDLAWLKVKDI-----PCSSALLTIDDNFCGLDMNAPLGVSDMVRGIPVF 425
Db 331 PFKEQETMNSNWLAVPSLKVPEPRPGQCVNDSRTLDP-----VSVNFVKSHTLMDAVPAF 386
QY 426 TEDDRMTSVIAVY-----KNHSLAFVGTSGKGLKXI-----RVD 461
Db 387 FTRPILIRISLQYRFTKIAVDQOVTPDQKAYDVLFIGTDDGKVIKALNSASFDSSDTVD 446
QY 462 GPRGNALQVETVQVDPGVLKDM---AFSKDHEQLYIMSERQLTRVPVYESCG--QYQSC 516
Db 447 -----SVVIEELQVLPFGVPVKNLYVVRMDGDSKLVVVSDDDEILAIKLRHCGSKITNC 501
QY 517 GECLSGDPHCGWCVLHNTCTRKERCERSKEPRRF-----ASEMKQC 558
Db 502 RECVSLODPYCAWDNVELKCTAVGSPDWSAGKRRFTQNTISLGEHKAC 548
```

Search completed: June 24, 2003, 19:01:28  
Job time : 38 secs





PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 170  
LENGTH: 552  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-905-291A-170

Query Match 13.58; Score 1351.5; DB 9; Length 552;  
Best Local Similarity 51.6%; Pred. No. 1.2e-97;  
Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;  
23 STLTTPAPLSPKOR-----SFTVTFGEPAE- 49  
45 SRLLT--AAPLSMEQRPWRALEVDSESVLLSVVWVLLAPPAAGMPQSFTHSENRDW 102  
50 GFNHLVVDERTGHIYLGAVNRIYKLSDLKVLVTHETGPDNDPKYPPRIQVTCNBLPT 109  
103 TFNHLTVHQGTGAVYVGAIRNRYKLTGATQVAHKTGPDNDKSRYPPLIVQPCSEVLT 162  
110 TTNVNVKLLIDYKENPLIACGSLYQICLKLRLEDLFLKGEVPHKHEHYLSGVNESGSV 169  
163 LTNVNVKLLIDYSENRLACGSLYQVGVCKLLRLDDFLVPEVSHKHEHYLSVNVKGTGM 222  
170 FGVIVSYSNLDDKLFATADVGPPEYPTTSSRKLTKNSADGMAFYVDFHDEFVASMKI 229  
223 YGVIVRSEGEDGKLFICTAVDGRQDYFTLSSRLPRDPRESSAMLDYELHSDFFVSLIKI 282  
230 PSTFTIIPDIYVYVGFSGNPFVFLTLQPMVSPG---STTKSQVYTSKVLRLCKE 286  
283 PSTDLAVSHFDIFYIYGFASGGFVYELTVQPE--TPEGVAINSAGDLFTYTSRIVRLCKD 340  
287 DTAFNSVVEPIGERSGVVELLOQAYLSKAGAVLGRTLGVHDDLLFTVFSKGGOKRK 346  
341 DPKFHSVSLPFGCTRAGVYRLQAYLAKPGDLSAQAFNITSQDDVLFVFAIFSGKQRY 400  
347 MKSLDESALCFTILKQINDRIKERLQSCYRGEGLDLAWLKVKDIPGSSALLTIDDDNFCG 406  
401 HHPDSDSALCAFPRAINLQIKERLQSCYQEGNLENNLLWLGKDVQCTKAPVPIDDDNFCG 460  
407 LDMNAPLGVSMDVRGIVFTEDRDRMTSVIAYVYKNSHSLAPVGTYSKGLKIRVDGPRGN 466  
461 LDINQPLGGSTPVEGLTLYTSRDRMTSVASYVYVYGVSVFVGTGSKGLKVRV----- 514  
467 ALOVE-----TVQVVDPGPVLRDMAFSK-DHEQLYIMSERQ 501  
515 ---YFRCNAIHLLESKESLLEGSYWWFRNTRQYFLGEQR 552

RESULT 2

US-09-902-853-170  
Sequence 170, Application US/09902853  
Publication No. US20020192659A1  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/902,853  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: US/09/665,350  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 170

; LENGTH: 552

; TYPE: PRT

; ORGANISM: Homo Sapien

US-09-902-853-170

Query Match 13.5%; Score 1351.5; DB 9; Length 552;

Best Local Similarity 51.6%; Pred. No. 1.2e-97;

Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;

```
QY 23 STLLTRQAPLPSOKOR-----SFTVTRGEPAE- 49
DB 45 SRLLT--AAPLSMEORQWPRALEVDSRSVLLSVVLLAPPAGMPQFSTFHSNRDW 102
QY 50 GFNHLVVDERTGHIYLGAVNRIYKLSDDLKVLVTHETGPDENPKCYPPIRVQTCNEPLT 109
DB 103 TFNHLTVHOGTGAIVYGAINRVYKLTGNLTQVAHKTGPEEDNKSRYPLIVQPCSEVLT. 162
QY 110 TTNVNVKMLLDYKENRLIACSLYOGICKLLRLDLFKLGEYPHKHEHYLSGVNESGSV 169
DB 163 LTNVNVKMLLDYSENRLIACSLYQGVCKLLRLDLFLVPSHKHEHYLSVNVKGTGM 222
QY 170 FGIVSVSNLDDKLFATAVDGKPEYFPTISRKLTKNSEADGMFAVYVHDEFVASMIKI 229
DB 223 YGVIVRSEGEDGKLFICTAVDGKQDYFPTLSRKLPRDPRESSAMLDYELHSDFFVSSLIKI 282
QY 230 PSDTFTIIPDFDIYVYGFSSGNFVYFLTLQPEMVSPPG---STTKEQVYVTSKLVRLCKE 286
DB 283 PSDTLALVSHFDIYVYGFSSGNFVYFLTLQPE---TPEGVAINSAGDLFTTSRIVRLCKD 340
QY 287 DPAFNSYVEVPTGCBRSYVEYRLLQAAYLSKAGAVLGRITLGVHPDDDLFTVPSKQKRR 346
DB 341 DPKFHSYVSLPGCTRAGVEYRLLQAAYLSKAGAVLGRITLGVHPDDDLFTVPSKQKRR 400
QY 347 MKSLDESALCIIFILQINDRIKERLOSRYGEGTLDLAWLKVDIPCSALLTIDNFCG 406
DB 401 HHPDDLSALCAPPAIRAINLQIKERLOSRYGEGTLDLAWLKVDIPCSALLTIDNFCG 460
QY 407 LDMNAPLGVSDVMRGIPVPTEDRDMTSVIAYVYKXHSFLAFVTKSGKLLKRVDDPRGN 466
DB 461 LDINQPLGSGTVEGLTLYTSDRMTSVASVYNGYSVVFVTKSGKLLKRVV-----514
QY 467 ALQYE-----TVQVNDPGPVLRDMAFSK-DHEQLYIMSERQ 501
DB 515 ---YEFRCNSATHLLSKESLLEGSYWMRFRNYRQLYFLGEOR 552
```

## RESULT 3

US-09-907-824-170

; Sequence 170, Application US/09907824

; Publication No. US20020197671A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

```
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT FILING DATE: 2001-07-17
; CURRENT APPLICATION NUMBER: US/09/907,824
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 170
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-907-824-170
```

Query Match 13.5%; Score 1351.5; DB 9; Length 552;

Best Local Similarity 51.6%; Pred. No. 1.2e-97;

Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;

```
QY 23 STLLTRQAPLPSOKOR-----SFTVTRGEPAE- 49
DB 45 SRLLT--AAPLSMEORQWPRALEVDSRSVLLSVVLLAPPAGMPQFSTFHSNRDW 102
QY 50 GFNHLVVDERTGHIYLGAVNRIYKLSDDLKVLVTHETGPDENPKCYPPIRVQTCNEPLT 109
DB 103 TFNHLTVHOGTGAIVYGAINRVYKLTGNLTQVAHKTGPEEDNKSRYPLIVQPCSEVLT 162
QY 110 TTNVNVKMLLDYKENRLIACSLYOGICKLLRLDLFKLGEYPHKHEHYLSGVNESGSV 169
DB 163 LTNVNVKMLLDYSENRLIACSLYQGVCKLLRLDLFLVPSHKHEHYLSVNVKGTGM 222
QY 170 FGIVSVSNLDDKLFATAVDGKPEYFPTISRKLTKNSEADGMFAVYVHDEFVASMIKI 229
DB 223 YGVIVRSEGEDGKLFICTAVDGKQDYFPTLSRKLPRDPRESSAMLDYELHSDFFVSSLIKI 282
QY 230 PSDTFTIIPDFDIYVYGFSSGNFVYFLTLQPEMVSPPG---STTKEQVYVTSKLVRLCKE 286
```

283 PSDTLALVSHFDIYIYGFASGGVFLTVQPE--TPEGVAINSAGDLFTYSTRIVRLCKD 340  
287 DTAFNSYVEVPIGERSGVYRLLQAAIYLSKAGAVLGRITLGVHPDDLLFTVTSKGOKRK 346  
341 DPKFHSYVSLPFGCTRAGVEYRLLQAAIYLSKAGAVLGRITLGVHPDDLLFTVTSKGOKQY 400  
347 MKSLDESALCIFILKQINDRIKERLQSCYRGGTDLAWLKVDIPCSSALLTIDNFCG 406  
401 HHPDDSALCAFPPIRINLQIKERLQSCYRGGTDLAWLKVDIPCSSALLTIDNFCG 460  
407 LDMNAPLGVSDMVRGIPVFTEDRDMTSAIVYVKNHSLAFVGTGSKLKIIRVDGPRGN 466  
461 LDINQPLGGSTPVEGLTLYTTSRDMTSAIVYVKNHSLAFVGTGSKLKIIRVDGPRGN 514  
467 ALQYE-----TVQVVDGPGVLRDMAFSK-DHEQLYIMSERQ 501  
515 ---YEFRCNAIHLLESKESLLEGSYWRFRNYRQLYFLGEOR 552

RESULT 4

US-09-907-841-170  
; Sequence 170, Application US/09907841  
; Publication No. US20020198366A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/907,841  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 170  
; LENGTH: 552  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-907-841-170  
  
Query Match 13.5%; Score 1351.5; DB 9; Length 552;  
Best Local Similarity 51.6%; Pred. No. 1.2e-97;  
Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;  
  
Qy 23 STLLTROPAPLSQKOR-----SFVTFRGEPAE- 49  
Db 45 SRLLT--AAPLSMEQRPWRALEVDSSVLLSVVVLLAPPAGMPQSFHSENROW 102  
Qy 50 GFNHLVVDERTGHIYLGAVNRIYKLSLDKLVIVTHTGPDENPKYPPRIVOTCNEPLT 109  
Db 103 TFNHLTVHQGTGAVYVGAIVNRYKLTGNLTITQVAHKTGPEENKSGRYPLIVQPCSEVLT 162  
Qy 110 TTNVNMKLLIDYKENRLIACGSLYQIGICKLLRLDLFKLGEYHKKHLYSLGVNESGSV 169  
Db 163 LTNNVNMKLLIDYSENRLIACGSLYQIGICKLLRLDLFKLGEYHKKHLYSLGVNESGSV 222  
Qy 170 FGVIVSVSNLDDKLFATVDGKPEYFPTISRKLTKNSEADGMPAYVHDFEFVSMIKI 229  
Db 223 YGVIVRSSEGDGKLFITGAVDGKQYFPTISRKLTKNSEADGMPAYVHDFEFVSMIKI 282  
Qy 230 PSDTFTIIPDEIYVYVGVSSGNFVYVLTLOPEMVSPG---STTKEQVYVTSKLVRLCKE 286  
Db 283 PSDTLALVSHFDIYIYGFASGGVFLTVQPE--TPEGVAINSAGDLFTYSTRIVRLCKD 340  
Qy 287 DTAFNSYVEVPIGERSGVYRLLQAAIYLSKAGAVLGRITLGVHPDDLLFTVTSKGOKRK 346  
Db 341 DPKFHSYVSLPFGCTRAGVEYRLLQAAIYLSKAGAVLGRITLGVHPDDLLFTVTSKGOKQY 400  
Qy 347 MKSLDESALCIFILKQINDRIKERLQSCYRGGTDLAWLKVDIPCSSALLTIDNFCG 406  
Db 401 HHPDDSALCAFPPIRINLQIKERLQSCYRGGTDLAWLKVDIPCSSALLTIDNFCG 460  
Qy 407 LDMNAPLGVSDMVRGIPVFTEDRDMTSAIVYVKNHSLAFVGTGSKLKIIRVDGPRGN 466  
Db 461 LDINQPLGGSTPVEGLTLYTTSRDMTSAIVYVKNHSLAFVGTGSKLKIIRVDGPRGN 514  
Qy 467 ALQYE-----TVQVVDGPGVLRDMAFSK-DHEQLYIMSERQ 501  
Db 515 ---YEFRCNAIHLLESKESLLEGSYWRFRNYRQLYFLGEOR 552

RESULT 5

US-09-904-011-170  
; Sequence 170, Application US/09904011  
; Publication No. US20030003530A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Mather, Jennie P.

APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/904,011  
CURRENT FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: 09/665,350  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 170  
LENGTH: 552  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-904-011-170

Query Match  
Best Local Similarity 13.5%; Score 1351.5; DB 9; Length 552;  
Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;  
SEQ ID NO 170

23 STLLTRQAPLKSQKQ-----SFVTPRGEPAE- 49  
45 SRLLT--AAPLSMEQRPWPRALEVDSRSVLLSVVWVLLAPPAAGMPQSFTHSENRDW 102  
50 GFNHLVVDERTGHVILGAVNRYKLSSDLKVLVTHETGDEDNPKYPPRIYVQTCNEPLT 109  
103 TFNHLTVHGTGAVVYGAINRVYKLTGNTIQVAKTGPEEDNKSRYPLIVQPCSEVLT 162  
110 TTNVNNKLLIYDYNKRLIACGSLYQIGICKLRLDLFKLGEYHKHKEHLSGWSGVS 169  
163 LTNVNNKLLIYDYNKRLIACGSLYQIGVCKLLRLDLFLVPESHKKEHLSVNTGTW 222  
170 FGIVSYSNLDDKLFATVADGKPEYFPTTSSRKLTKNSEADGMFAIVPHDEFVSMIKI 229

Db 223 YGVIVRSEGDGKLFICTAVDGKQDYPTTSSRKLPRDPRESSAMLDYELHSDFFVSSLIKI 282  
QY 230 PSDTFTIIPDFDIYVYVGFSSGNFVYFELTLOPEMVSPG---STTKEQVYTSKLVRLCKE 286  
Db 283 PSDTLALVSHFDIFYIYVGFASGFFVYFELTVQPE--TEGVAINAGDDLFTYSRIVRLCKD 340  
QY 287 DTAFNSYVEVPIGERSGVYRLLQAAAYLSKAGAVLQRTLGVHPDDDLFTVFSKGQRK 346  
Db 341 DPRFHSYVSLPFGCTRAGVEYRLLQAAAYLAKPGDSLAAQAFNITSQDDVLFVFAFSKGQKQY 400  
QY 347 MKSLDESALCIPILKQINDRIKERLQSCYREGTGLDLAWLKVKDIPCSSALLTIDDNFCG 406  
Db 401 HPPDDDSALCAFFIRAINLQIKERLQSCYQEGNLELNWLLGKDVQCTKAPVPIDDNFCG 460  
QY 407 LDMNAPLGVSMDVIRGIPVFTEDRDMTSTVYAYVYKHSIAFVGTGSKGLKKIRVDGPRGN 466  
Db 461 LDINQPLGGSTPVEGLTYTTSRDMTSTVASYVYNGYSVVFVGTGSKGLKKVRV----- 514  
QY 467 ALQYE-----TVQVVDPPVLRDMAFSK-DHEQLYIMSERQ 501  
Db 515 ---YFRCNSAIHLLSKESLESLSYWMRFVYRQLYFLGEQR 552

RESULT 6  
US-10-028-072-332  
; Sequence 332, Application US/10028072  
; Publication No. US20030004311A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang  
; TITLE OF INVENTION:  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/028,072  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059184  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059352  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059836  
; PRIOR FILING DATE: 1997-09-24  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



```
QY 230 PSTFTIIPDFIYVYVSSGNFVFLTLQPEMVSPG---STTKQVYTSKLVRLCKE 286
DB 283 PSTDLALVSHFDIYIYGFASGGFVFLTVQPE--TPEGVAINSAGDLFTYSRIVRLCKD 340
QY 287 DTAFNSVVEPIGCSRGVEYRLLQAAAYLSKAGAVLGRITLGVHPDDDLLFTVFSKGQKRR 346
DB 341 DPKFHSVSLPFGCTRAGVEYRLLQAAAYLAKPGDSLAAQAFNITSQDDVLFALFSGKQKQY 400
QY 347 MKSLDESALCIFIILKQINDRIKERLQSCYRGEGTLDLAWLKVKDIPCSSALLTIDDNFCG 406
DB 401 HHPDDLSALCAFPPIRALNLOIKERLQSCYQEGNLELNLGDKVQCTKAPVPIDDNFCG 460
QY 407 LDMNAPLGSDMVRGIPVFTEDRDMTSLVAYVYKNHSLAFVGTGSKGLKKIRVDGPRGN 466
DB 461 LDINQPLGGSTPVEGLTLYTTSRDMTSSVSVYNGYSVVFVGTGSKGLKKVRV-----514
QY 467 ALOYE-----TVQVVDPGVLRDMAFSK-DHEQLYIMSERQ 501
DB 515 ---YEFRCNAIHLLESKESLLEGSYWWRFNRQLYFLGEQR 552

RESULT 8
US-10-121-049-332
; Sequence 332, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 332
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-332

Query Match 13.5%; Score 1351.5; DB 9; Length 552;
Best Local Similarity 51.6%; Pred. No. 1.2e-97;
Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;

QY 23 STLLTRQAPLPSOKOR-----SFTVFRGEPAE- 49
DB 45 SLLLT--AAPLSMEQRPWRALEVDRSVLLSVVWLLAPPAGMPQFSTFHSNRDW 102
QY 50 GFNHLVVDERTGHIYLGAVNRIYKLSDDLKVLVTHETGPDENPKCYPPRIVQTCNEPLT 109
DB 103 TFNHLTVHOGTGAVYVGAIRNVKLTGNLTIOVAHKTGPEEDNKSRYPLLIQPCSEVL 162
QY 110 TTNVNMKLLIDYKENRLIACGSLYQGI CKLRLLEDLFKLGEPYHKKEHYLSGVNCSGV 169
DB 163 LTNVNMKLLI IDYSENRLLACGSLYGVCKLLRLDLDLILVEPSHKKEHYLSVNVKTGM 222
QY 170 FGVIYSNLDKLFATATVDGKPEVFTTSSRKTNSKNSADMGFAVVFHDEFVASMIKI 229
```

```
DB 223 YGVIVRSEGEDGKLFICTAVDQKDYFPTLSSRKLPRDPRESSAMLDYELHSDFVSSLIKI 282
QY 230 PSTFTIIPDFIYVYVSSGNFVFLTLQPEMVSPG---STTKQVYTSKLVRLCKE 286
DB 283 PSTDLALVSHFDIYIYGFASGGFVFLTVQPE--TPEGVAINSAGDLFTYSRIVRLCKD 340
QY 287 DTAFNSVVEPIGCSRGVEYRLLQAAAYLSKAGAVLGRITLGVHPDDDLLFTVFSKGQKRR 346
DB 341 DPKFHSVSLPFGCTRAGVEYRLLQAAAYLAKPGDSLAAQAFNITSQDDVLFALFSGKQKQY 400
QY 347 MKSLDESALCIFIILKQINDRIKERLQSCYRGEGTLDLAWLKVKDIPCSSALLTIDDNFCG 406
DB 401 HHPDDLSALCAFPPIRALNLOIKERLQSCYQEGNLELNLGDKVQCTKAPVPIDDNFCG 460
QY 407 LDMNAPLGSDMVRGIPVFTEDRDMTSLVAYVYKNHSLAFVGTGSKGLKKIRVDGPRGN 466
DB 461 LDINQPLGGSTPVEGLTLYTTSRDMTSSVSVYNGYSVVFVGTGSKGLKKVRV-----514
QY 467 ALOYE-----TVQVVDPGVLRDMAFSK-DHEQLYIMSERQ 501
DB 515 ---YEFRCNAIHLLESKESLLEGSYWWRFNRQLYFLGEQR 552

RESULT 9
US-10-123-904-332
; Sequence 332, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 332
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-904-332

Query Match 13.5%; Score 1351.5; DB 9; Length 552;
Best Local Similarity 51.6%; Pred. No. 1.2e-97;
Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;

QY 23 STLLTRQAPLPSOKOR-----SFTVFRGEPAE- 49
DB 45 SLLLT--AAPLSMEQRPWRALEVDRSVLLSVVWLLAPPAGMPQFSTFHSNRDW 102
QY 50 GFNHLVVDERTGHIYLGAVNRIYKLSDDLKVLVTHETGPDENPKCYPPRIVQTCNEPLT 109
DB 103 TFNHLTVHOGTGAVYVGAIRNVKLTGNLTIOVAHKTGPEEDNKSRYPLLIQPCSEVL 162
QY 110 TTNVNMKLLIDYKENRLIACGSLYQGI CKLRLLEDLFKLGEPYHKKEHYLSGVNCSGV 169
```



1 PRIOR APPLICATION NUMBER: US 60/146,222  
2 PRIOR FILING DATE: 1999-07-28  
3 PRIOR APPLICATION NUMBER: PCT/US99/20594  
4 PRIOR FILING DATE: 1999-09-08  
5 PRIOR APPLICATION NUMBER: PCT/US99/20944  
6 PRIOR FILING DATE: 1999-09-13  
7 PRIOR APPLICATION NUMBER: PCT/US99/21090  
8 PRIOR FILING DATE: 1999-09-15  
9 PRIOR APPLICATION NUMBER: PCT/US99/21547  
10 PRIOR FILING DATE: 1999-09-15  
11 PRIOR APPLICATION NUMBER: PCT/US99/23089  
12 PRIOR FILING DATE: 1999-10-05  
13 PRIOR APPLICATION NUMBER: PCT/US99/28214  
14 PRIOR FILING DATE: 1999-11-29  
15 PRIOR APPLICATION NUMBER: PCT/US99/28313  
16 PRIOR FILING DATE: 1999-11-30  
17 PRIOR APPLICATION NUMBER: PCT/US99/28564  
18 PRIOR FILING DATE: 1999-12-02  
19 PRIOR APPLICATION NUMBER: PCT/US99/28565  
20 PRIOR FILING DATE: 1999-12-02  
21 PRIOR APPLICATION NUMBER: PCT/US99/30095  
22 PRIOR FILING DATE: 1999-12-16  
23 PRIOR APPLICATION NUMBER: PCT/US99/30911  
24 PRIOR FILING DATE: 1999-12-20  
25 PRIOR APPLICATION NUMBER: PCT/US99/30999  
26 PRIOR FILING DATE: 1999-12-20  
27 PRIOR APPLICATION NUMBER: PCT/US00/00219  
28 PRIOR FILING DATE: 2000-01-05  
29 NUMBER OF SEQ ID NOS: 423  
30 SEQ ID NO 170  
31 LENGTH: 552  
32 TYPE: PRT  
33 ORGANISM: Homo Sapien  
34 US-09-906-838-170

Query Match 13.5%; Score 1351.5; DB 9; Length 552;  
Best Local Similarity 51.6%; Pred. No. 1.2e-97;  
Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;

QY 23 STLLTRQAPLSQOKR-----SFTVTFGEPAE- 49  
DB 45 SRLLT--AAPLSMEQRPWPALEVDGRSVLLSVVWVLLAPPAGMQPQFSTFISENRDW 102  
QY 50 GFNLVVDERTGTHYLGAIVNRNYKSLSDKLKVLVTHETGPDENPKCYPPRIQVTCNEPLT 109  
DB 103 TFNHLTVHQGTGAVYGAINRVKLTGNLTITQVAHKTGPEEDNKSRYPLIVQPCSEVLT 162  
QY 110 TTNVNVKMLLDYKENRLIACGSLYQIGICKLLRLLEDLPKLGEPYHKKHLYSGVNESGSV 169  
DB 163 LTNVNVKMLLDYKENRLIACGSLYQIGICKLLRLLEDLPKLGEPYHKKHLYSGVNESGSV 222  
QY 170 FGTVSYSNLDDKLFATADVGRPEYPTTISRSKLTKNSEADGFMFAVVFHDEFVASMIKI 229  
DB 223 YGVIVRSEGEDGKLFITAVDGKQDYPTLTSRKLPRDPESAMLDYELHSDFFVSSLIKI 282  
QY 230 PSDTFTIIPDIIYVYVGFSSGNFYFLTLQPEVMSPEG---STTKQVYTSKLVRLCKE 286  
DB 283 PSDTLALVSHFDIYVYVGFSSGNFYFLTLQPEVMSPEG---STTKQVYTSKLVRLCKD 340  
QY 287 DTANSVYVEVIGERSGVETRLQAAVLSKAGAVLGTGLVHPDDDLTFVFSKGOKRK 346  
DB 341 DPKFHSYVSLPFGCTRAGVEYRLQAAVLSKAGAVLGTGLVHPDDDLTFVFSKGOKRK 400  
QY 347 MKSLDEGALCFILKQINDRIKERLQSCYRCEGTLDLAWLKVDIIPCSSALLTIDDNFCG 406  
DB 401 HHPDDGALCAFPRAINLQIKERLQSCYRCEGTLDLAWLKVDIIPCSSALLTIDDNFCG 460  
QY 407 LDMNAPLGVSMDVRGIPVFTEDRDMTSSVIAVYVYKNSLAFVGTGSKGLKIRVDGPRGN 466  
DB 461 LDINQPLGGSTPVEGLTLYTTSRDGMTSSVIAVYVYKNSLAFVGTGSKGLKIRVDGPRGN 514  
QY 467 ALQYE-----TVQVVDPCPLVLRDMAFSK-DHEQLYIMSERQ 501

Db 515 ---YEFRCRNAIHLLSKESLLEGSYWMRFNRYRQLYFLGEQR 552

RESULT 12  
US-09-907-613-170  
; Sequence 170, Application US/09907613  
; Publication No. US20030027145A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Aeshkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kijavir, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/907,613  
; CURRENT FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423

SEQ ID NO 170  
LENGTH: 552  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-907-613-170

Query Match 13.5%; Score 1351.5; DB 9; Length 552;  
Best Local Similarity 51.6%; Pred. No. 1.2e-97;  
Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;

QY 23 STLLTRQAPLSQKOR-----SFTVTRGEPAE- 49  
DB 45 SKLLT--AAPLSMEQRPWRALEVDSSRVLLSVVLLAPPAGMPQFSTFHSNRDM 102  
QY 50 GFNHLVVDERTGHIYLGAVNRYYKLSDDLKVLVTHETGPDENPKCYPRIVQTCNEPLT 109  
DB 103 TFNHLTVHQGTGAVYGAINRVYKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLT 162  
QY 110 TTNVNNKMLLDYKENRLIACGSLYQGVCKLRLDLDLFLVPEPSHKKEHYLSVNVKTGT 169  
DB 163 LTNNVNNKMLLDYSENRLLACGSLYQGVCKLRLDLDLFLVPEPSHKKEHYLSVNVKTGT 222  
QY 170 FGIVYSYNLDDKLFATADCKPEYFPTISSRKLTKNSEADGMFAYVPHDFVASMIMI 229  
DB 223 XGVIVRSEGEDGKLFICTAVDGKQDYFPTLSSRKUPRDPRESSAMLDYELHSDFFVSSLIKI 282  
QY 230 PSDTFIIPDFDIYVYGFSSGNFVYFLTLQPEWSPPG---STTKEQVYTKLVRCKE 286  
DB 283 PSDTLALVSHFDIFYIYGFASGGFVFLTVQPE--TPEGVAINSAGDLFTTSRIVRLCKD 340  
QY 287 DPAFNSYVEVPICGRSGVEYELQAAQLYLSKAGAVLGRITLGVHPDDDLFTVFSKGQK 346  
DB 341 DPKFSYVSLPGCTRAGVEYELQAAQLYLSKAGAVLGRITLGVHPDDDLFTVFSKGQK 400  
QY 347 MKSLDESALCIFILQINDRIKRIERLOSRYGSGTDLNWKDIPCSALITIDNFCG 406  
DB 401 HHPDDLSALCAPPAIRAINLOIKERLQSCYQSGNLELNLGKDVQCTKAPVIDDNCFCG 460  
QY 407 LDMAPLGVSDMVRGIPVETEDRMTSVIAVYKXHSIAFVGTSGKLRKIRVODGPRGN 466  
DB 461 LDINPLGSGTVEGLTYTTSRDMTSSVSYNGYGVVFGTSGKLRKIRV-----514  
QY 467 ALQYE-----TVQVVDGPGVLRDMAFSK-DHRLQYIMSERQ 501  
DB 515 ---YEFRCNAITHLSKESLLEGSYVWRNRYRQLYFLGEOR 552

## RESULT 13

US-09-907-942-170  
Sequence 170, Application US/09907942  
Publication No. US20030027146A1

## GENERAL INFORMATION:

APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertitsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT FILING DATE: 2002-01-22  
CURRENT APPLICATION NUMBER: US/09/907,942  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 170  
LENGTH: 552  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-907-942-170

Query Match 13.5%; Score 1351.5; DB 9; Length 552;  
Best Local Similarity 51.6%; Pred. No. 1.2e-97;

Matches 269; Conservative 81; Mismatches 116; Indels 55; Gaps 8;

QY 23 STLLTRQAPLSQKOR-----SFTVTRGEPAE- 49  
DB 45 SKLLT--AAPLSMEQRPWRALEVDSSRVLLSVVLLAPPAGMPQFSTFHSNRDM 102  
QY 50 GFNHLVVDERTGHIYLGAVNRYYKLSDDLKVLVTHETGPDENPKCYPRIVQTCNEPLT 109  
DB 103 TFNHLTVHQGTGAVYGAINRVYKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLT 162  
QY 110 TTNVNNKMLLDYKENRLIACGSLYQGVCKLRLDLDLFLVPEPSHKKEHYLSVNVSGSV 169  
DB 163 LTNNVNNKMLLDYSENRLLACGSLYQGVCKLRLDLDLFLVPEPSHKKEHYLSVNVKTGT 222  
QY 170 FGIVYSYNLDDKLFATADCKPEYFPTISSRKLTKNSEADGMFAYVPHDFVASMIMI 229  
DB 223 XGVIVRSEGEDGKLFICTAVDGKQDYFPTLSSRKUPRDPRESSAMLDYELHSDFFVSSLIKI 282  
QY 230 PSDTFIIPDFDIYVYGFSSGNFVFLTLQPEWSPPG---STTKEQVYTKLVRCKE 286  
DB 283 PSDTLALVSHFDIFYIYGFASGGFVFLTVQPE--TPEGVAINSAGDLFTTSRIVRLCKD 340



Db 223 YGVIIRSEGEDKLFITGTAVDGKQDYFPTLSSRKLPDPPESSAMLDYELHSDFVSSLIKI 282  
Qy 230 PSDTFIIPDFIYVYGFSSGNFVYFLTLQPEMWSPPG---STTKEQVYVTSKLVRLCKE 286  
Db 283 PSDTLALVSHFDIFYIGFASGFFVFLTVQPE--TPEGVAINSGDLFTYSRIVRLCKD 340  
Qy 287 DPAFNSYVEVPIGCBERSGYEYLLQAAYLKAGAVLGRITLGVHPDDDLLFTVFSKGQKRX 346  
Db 341 DPKFHSYVSLPFGCTRAGVEYRLLQAAVLAKPGDSLQAQFNITSQDDVLFALFSKGQKQY 400  
Qy 347 MKSLDESALCIFIILKQINDRIKRIQSCYRGEGTLDLAWLKVDIPCSSALLTIDNFCG 406  
Db 401 HHPPDDLSALCAFPPIRAINLQIKERLQSCYQEGNLELNLWLLGKDVQCTKAPVIDNFCG 460  
Qy 407 LDMNAPLGVS DMVRGIPVFTEDRDMTISVIAVYKXHSIAFVGTSGKGLKIRVDGPRGN 466  
Db 461 LDINQPLGGSTPVEGLTYTTSRDMTISVASYVYNGYSVVFVGTSGKGLKIRV----- 514  
Qy 467 ALQYE-----TVQVVDPGVLRDMAFSK-DHEQLYIMSERQ 501  
Db 515 ---YEFRCNSAITHLLSKESLLGSGYWWRFNYRQLYFLGQOR 552

Search completed: June 24, 2003, 19:10:54  
Job time : 73 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 18:54:06 ; Search time 62 Seconds  
(without alignments)  
2939.853 Million cell updates/sec

Title: US-09-964-956-13  
Perfect score: 9990  
Sequence: 1 MKAMPWNWTCLLSHLLVMGM.....QKLAYKLEQVITMSLDSNK 1896  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6764	67.7	1884	JC4975	plexin 2 precursor
2	6393.5	64.0	1905	I51553	Plexin - African c
3	6337	63.4	1894	JC4980	plexin 1 precursor
4	5984.5	59.9	1872	JC4976	plexin 3 precursor
5	3631	36.3	1945	T13937	plexin A - fruit f
6	2932.5	29.4	2051	T13164	plexin B - fruit f
7	1609.5	16.1	1806	T23298	hypothetical prote
8	1349	13.5	1568	T09074	semaphorin recepto
9	1322	13.2	317	T46426	hypothetical prote
10	766	7.7	177	T47133	hypothetical prote
11	652	6.5	1375	JCS148	hepatocyte growth
12	651	6.5	1390	TVHUME	hepatocyte growth
13	646.5	6.5	1379	S01254	hepatocyte growth
14	559	5.6	1375	T30813	plasmaogen relate
15	556	5.6	1425	T30811	hepatocyte growth
16	538.5	5.4	1404	A48196	protein-tyrosine k
17	522	5.2	349	I52685	gene MM1 protein -
18	514.5	5.2	1369	JC4860	protein-tyrosine k
19	480.5	4.8	1374	T30809	plasmaogen relate
20	427	4.3	1400	I38185	protein-tyrosine k
21	387.5	3.9	1378	I48751	protein-tyrosine k
22	308.5	3.1	782	I48746	semaphorin C - mou
23	292.5	2.9	1074	JCS928	semaphorin F precu
24	244.5	2.4	760	I48745	semaphorin B - mou
25	227	2.3	712	T27165	hypothetical prote
26	227	2.3	724	C49423	semaphorin II prec
27	207	2.1	730	JH0798	fasciclin IV precu
28	205.5	2.1	834	S66498	M-sema F protein p
29	203	2.0	772	I48747	semaphorin D - mou

30	201	2.0	749	2	G01856	semaphorin V - hum
31	200	2.0	771	2	D49423	semaphorin III pre
32	197	2.0	748	2	I48744	semaphorin A - mou
33	186.5	1.9	711	2	A49423	semaphorin I precu
34	185.5	1.9	676	2	T33853	hypothetical prote
35	179.5	1.8	656	2	B49423	semaphorin I - fru
36	177.5	1.8	753	2	G02173	semaphorin III fam
37	176	1.8	772	2	A49069	collapsin - chicke
38	170	1.7	751	2	I48748	semaphorin E - mou
39	163.5	1.6	666	2	I58169	semaphorin III - m
40	162	1.6	642	2	T29757	protein UNC-89 - C
41	154	1.5	3856	2	T51174	ataxia-telangietta
42	147.5	1.5	4307	2	T20721	hypothetical prote
43	146.5	1.5	653	2	T03102	semaphorin homolog
44	144.5	1.4	5175	2	T20992	hypothetical prote
45	144.5	1.4	5198	2	T43290	hemicentin precurs

ALIGNMENTS

RESULT 1

JC4975  
plexin 2 precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 05-Nov-1999  
C;Accession: JC4975  
R;Kameyama, T.; Murakami, Y.; Suto, F.; Kawakami, A.; Takagi, S.; Hirata, T.; Fujisawa, T.  
Biochem. Biophys. Res. Commun. 226, 396-402, 1996  
A;Title: Identification of plexin family molecules in mice.  
A;Reference number: JC4975; MUID:96400270; PMID:8806646  
A;Accession: JC4975  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-1884 <AM>  
A;Cross-references: DDBJ:D86949; NID:gl555431; PIDN:BAA13189.1; PID:dl013878; PID:gl555431  
C;Comment: This protein is a membrane protein with cell adhesion properties.  
C;Keywords: duplication; transmembrane protein  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;501-549/Region: cysteine-rich duplication  
F;646-691/Region: cysteine-rich duplication  
F;794-845/Region: cysteine-rich duplication  
F;1223-1250/Domain: transmembrane #status predicted <TMM>

Query Match	67.7%	Score	6764;	DB	2;	Length	1884;		
Best Local Similarity	67.5%;	Pred. No.	0;						
Matches	1267;	Conservative	239;	Mismatches	356;	Indels	14;	Gaps	8;
QY	25	LLTRQAPAPLSQKQRSFVTPRGPAPAE-GFNHLVVDERTGHIYLGAVNRIYKLSDDLKVLVT	83						
DB	17	LLARGTTGMPQ---YSTFHSNRDWTFFNHLTVHRRTGAVTVGAINRVYKLTGNLTQVA	72						
QY	84	HETGPDNDPKYPPRIYVQCNEPLTTNNVNKMLIDYKENRLIACGSLYQGIKLLRL	143						
DB	73	HKTGPEEDNKACYPPLIVQPCSEVLTITNNVNKMLIDYSENRLIACGSLYQGIKLLRL	132						
QY	144	EDLFLKGLGEYHKKHEHLSVGNESGSVFGIVSYNSNLDLKLFIATVDGKPEYPTTSSRK	203						
DB	133	DOLFILVPEPSHKHEHLSVGNESGVNKTGMGYIVRSEGEDGKLFICTAVDVGKQDYPTTSSRK	192						
QY	204	LTKNSEADGMFAFYVDFHDFVASMIIKIPSDTFTIIPDFIYVYVGFSSGNFYFLTQPEM	263						
DB	193	LPRDPSSAMLDYELHSDFFVSSLIKIPSDTALVSHDFIFYIYGFASGGGFYFLTQPE-	251						
QY	264	VSPPG---STTKEQVYTSKLVRICKEDTAFNSVVEYPIGCGERSGVETRLLOAAYLSKAGA	320						
DB	252	-TPDGMAINSAGDLFYTSRIVRLCKDDPKFHSVYSLFPFGCTGAGVEYRLLLOAAYLAKPGE	310						
QY	321	VLGRTLGVHPDDLLFTVFSKGQRKMKSLDESALCIFIILKQINDRIKERLQSCYRGEGT	380						
DB	311	ALAAQAFNISDEDLVFAIFRSKGQRKMKSLDESALCIFIILKQINDRIKERLQSCYHGEGN	370						
QY	381	LDLAWLKVKDIPCSSALLTTIDNFCGLDMMNAPLGVSMDVMVRGIPVFYFEDRDMTSVIAYVY	440						

Db	371	LELNLGKVOCTKAPVPIDDDNECGLDINQPLGGSTPVEGLTYTTSRDLTSVASVY	430	Db	1449	FMLYCAIKQOMEKGPIDAITGEARYSLSEDKLIRQOIEYKTLILNCVNPDPNENSPPEIPVK	1500
Qy	441	KHSLAFVGTSGKLRKIRVDGPRGNALQYETVVV-DPGVLRDMARSKDHEOLYTMSE	499	Qy	1519	ILNCDDTITQVKEKILDAIFKNVPCSHRPAKADMDLEWRQSGARMILQDDEDITTKIENDW	1571
Db	431	NGYSVVFVGTSGKLRKIRADGPPHGGVQYEMVSFVKDGPILRDMASINQILYLYMSE	490	Db	1509	VLNCDDTITQVKEKILDAVYKNVPYSORPRAVMDLEWRQGIARVVLQDDEDITTKIEGDM	1566
Qy	500	QLTRVPVESCQYQSCGCGCLSGDPHCGMVLHNTCTRKCERCSEKPRRFASEMKCQC	559	Qy	1579	KRLNTLAHVQVDPGVSVALVSKQVTAAYNAVNSTVSRTSASKEYENMIRYTGSPDLSRST	1633
Db	491	RQVTRVPVESCQYITTCGECUSSGDPHCGMVLHNTCTRKCERCSEKPRRFASEMKCQC	550	Db	1569	KRLNTLAHVQVDRSVVALVPKQTSYNNIPASASISRTSISRYDSSFRYTGSPDLSRST	1626
Qy	560	RLTVHPNINISVSNVLLVETYNVPELSAGVNCPTFEDLSMDGLVVGNOIQCVSPRAKE	619	Qy	1639	PMITPDLESQVKNVHLVKNHEHGDQEGDQSGKVMSEIYLTLLATKGTLOKFVDDLPET	1663
Db	551	SLEVPNISTSVSDHSLLSLVNDAPNLSEGIACAFGNLTVEGQVSGSQVICISPGPKD	610	Db	1629	PMITPDLESQVKNVHLVKNHHDGQEGDQSGKVMSEIYLTLLATKGTLOKFVDDLPET	1663
Qy	620	VPRITENGHHVLOLQKSETGMTFASTSVFYNGSVHNSCLSCVESPYRCHWCKYRH	679	Qy	1699	IPSTAHRSALPLAIKYMDFDLDEQADKHGIDHPVHRTWKNCLPLRFFWVNMKNPQFV	1771
Db	611	VP-VIPLDQDFWGLEQLRSKETGKIFVSTEFKPYNCASAHQJCLSCVNSAFRCHWCKYRN	669	Db	1689	LFSTVHRGSLPLAIKYMDFDLDEQADKHGIDHPVHRTWKNCLPLRFFWVNMKNPQFV	1771
Qy	680	VCTHDPKTCFQGRVVKLPEDCPQLLRVDKILVPEVIKPIITLKAKNLPOQSGQGYEC	739	Qy	1759	FDIHKNSITDACLSSVAQTQFMDSCSTSEHRLGKQSPSNKLLYAKDIPSYKNVVERYSI	1818
Db	670	LCTHDPKTCFQGRVVKLPEDCPQLLRVDKILVPEVIKPIITLKAKNLPOQSGQGYEC	729	Db	1749	FDIHKNSITDACLSSVAQTQFMDSCSTSEHRLGKQSPSNKLLYAKDIPSYKNVVERYSI	1818
Qy	740	ILNIQSEORVPALRNFSSVQCNQTSYSYEGMEINNLPLVELTVVWNGHFNIDNPAQNV	799	Qy	1819	GHMPALSDQDMNAYLAEQSRMHNENFTMSALSEIFSYVGYKSEIILGDLHDDQCGKQK	1818
Db	730	VLSIQGAVHRVPALRNFSSVQCNQTSYSYEGMEINNLPLVELTVVWNGHFNIDNPAQNV	789	Db	1809	AKLPAISDQDMNAYLAEQSRMHNENFTMSALSEIFSYVGYKSEIILGDLHDDQCGKQK	1818
Qy	800	HLKCGAMRESGLCLKADDPACGWCQGGCOCTLRQHPAQOESQWLELGAKSCTNPR	859	Qy	1879	LAYKLEOVITILMSLDS 1894	
Db	790	HLKCAAAQRESGLCLKADDPACGWCQGGCOCTLRQHPAQOESQWLELGAKSCTNPR	849	Db	1869	LAYKLEOVITILMSLDS 1884	
Qy	860	ITEIIPVTGREGTKVIRGENLGLBFRDIASHVKVAGVECSPLVDGYIPAEQIVCEMG	919				
Db	850	ITEIIPVTGREGTKVIRGENLGLBFRDIASHVKVAGVECSPLVDGYIPAEQIVCEMG	909				
Qy	920	EAKPSQHAGFVEICVAVCEPEFPMARSSQLYFMTLTLSDLKPSRGPMSGGTQVITGTNL	979				
Db	910	HAVIGTSGPRLCTGECPEFMTKSHQOYTFVNFVLSLSPRGESGGTQVITGTNL	969				
Qy	980	NAGSNVVMFCQKPCFLHRSRPSYIVC-NTTSSDEVLEMKVSVQVDRAKIHODLVQYVE	1038				
Db	970	GAGSSVAVLGNQTCFEGYGRSMNEIVCVSPSSNGLGFPVSVSDRARDVSSLOPEYID	1029				
Qy	1039	DPTVRIPEWSIVSGNTPIAVGTHLIDLIQNPQIRAKHGKHEINICEVLNATEMTQCA	1098				
Db	1030	DPVQRIPEWSIVSGNTPIAVGTHLIDLIQNPQIRAKHGKHEINICEVLNATEMTQCA	1089				
Qy	1099	PALALGPDHQDLTERPEEFGFILDNVQSLILINKNTFTYYPNPVFEAFPGSGILELKPQ	1158				
Db	1090	PSLTSYRPLDTERPEEFGFILDNVQSLILINKNTFTYYPNPVFEAFPGSGILELKPQ	1149				
Qy	1159	TPILKGNLIPPVAGNVKLYTVLVGKPCPTVSDVQLLCESPNLIGRHKVMARVGG	1218				
Db	1150	SPILKGNLIPPVAGNVKLYTVLVGKPCPTVSDVQLLCESPNLIGRHKVMARVGG	1208				
Qy	1219	MEYSPGMVYIADPSLPAISVAVAGLLIIFIVAVLIAYKRKRESDLTLKRLQOM	1278				
Db	1209	MVFPSPGVSVIDSLTLPLAISVAVAGLLIIFIVAVLIAYKRKRESDLTLKRLQOM	1268				
Qy	1279	DNLSRVALECKEAFELQTDIHELTSLDGAGIFPDYRTVMVLPFGIEDHPVLRDL	1338				
Db	1269	DNLSRVALECKEAFELQTDIHELTSLDGAGIFPDYRTVMVLPFGIEDHPVLRDL	1328				
Qy	1339	EVPGVQERVEKGLKLAQNLNNKVPFLSIFRTLESORFSRMDRGNVASLIMTVLQSKL	1398				
Db	1329	EVQNGQOQVEKALKLAQNLNNKVPFLSIFRTLESORFSRMDRGNVASLIMTVLQSKL	1388				
Qy	1399	EYATDVLKOLLADLIDKNLESKNHPKLLRRTESVAEKMLTNWFTLLYFLKCECAGEPL	1458				
Db	1389	EYATDVLKOLLADLIDKNLESKNHPKLLRRTESVAEKMLTNWFTLLYFLKCECAGEPL	1448				
Qy	1459	FSLFCAIKQOMEKGPIDAITGEARYSLSEDKLIRQOIEYKTLILNCVNPDPNENSPPEIPVK	1518				

RESULT 2  
151553  
Plexin - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C;Accession: 151553  
R;Ohta, K.; Mizutani, A.; Kawakami, A.; Murakami, Y.; Kasuya, Y.; Takagi, S.  
Neuron 14, 1189-1199, 1995  
A;Title: Plexin: a novel neuronal cell surface molecule that mediates cell ad  
A;Reference number: 151553; MUID:95329274; PMID:7605632  
A;Accession: 151553  
A;Status: Preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1905 <OHT>  
A;Cross-references: GB:D38175; NID:961514; PIDN:BA07374.1; PID:961515  
C;Keywords: duplication

Qy	6	WNWTCLLSHLLMVGMGSSTLTTRQAPLQKQSRQSVFTFRGEPAGFNHVVDDERTGHIYL	6	Qy	126	RLIACGSYQGIQCKLRLLEDLFLKLGEPYHKEHYLSGVNESGVSFGVIVSYSNLDDKLF	1
Db	14	WTFVLVLLGSWATGDSG-----PKDFRTFTGSD-WSLTHLVVHNKTGEVIV	5	Db	119	RLIACGSASOGICQFLRLDDLFLKLGEPHHRKEHYLSGVNESGTMGSGVIEVPNGQNKLF	1
Qy	66	GAVNRIYKLLSSDLKVLVTHETGPDENPKCYPRIVQTCNEPLTTNNVKNKMLLDYKEN	1	Qy	186	ATAVDGKPEYFPTLSSRKLTKNSADGMAVFEHDEFVASMIKIPSDTFTTIIIPDFDIYV	2
Db	59	GAINRIYKLSNNLTLTHVTGVEDNEKCYPPSPVQSCPHGLITNNVKNKMLLDYSDN	1	Db	179	GTPIDGKSEYFPTLSSRKLKGNEENAEFMFVQDEFVSSQLKIPSDTSLSKFPTDIYV	2
Qy	126	RLIACGSYQGIQCKLRLLEDLFLKLGEPYHKEHYLSGVNESGVSFGVIVSYSNLDDKLF	1	Qy	246	YGFSSGNFVYFLTLQ--PEMVSPPGSTTKEQVYTSKLVLCKEDTAFNSYVEVPICGERS	3
Db	119	RLIACGSASOGICQFLRLDDLFLKLGEPHHRKEHYLSGVNESGTMGSGVIEVPNGQNKLF	1	Db	239	YGFSSQFVYVYTLTQLDQLTSP--DSTGEQFTTSKIVRLCVDDPKFYSEVFPICGMD	2
Qy	186	ATAVDGKPEYFPTLSSRKLTKNSADGMAVFEHDEFVASMIKIPSDTFTTIIIPDFDIYV	2	Qy	304	GVEYRLLQAAVLSKAGAVLGRTLGVHDDDDLFTVFSKQGRKMKSLDESALCIFILKQI	3

QY	1099	PALALGPDRHQSDLTERPEEFGFI	LDNVOSLLILNKNTFTYYPNPFEAF	PGSGILLEKPG	1158		
DB	1090	PSLTSDYREGLDVTVERPDEFGLF	FNNVOSLLIYNDTKFIYYPNPTFELL	SPGILDLQKPG	1149		
QY	1159	TPIIILKGNLI	PPVAGGNVKNLYTLVGSBKPC	TVTVSDVQLLCESPNLIGRHKVMARVGG	1218		
DB	1150	SPIIILKGNLCP	PASGG-AKUNYTVMI	GETPCVTYVSETQLLCEPPLTGQHKVMHVHVG	1208		
QY	1219	MEYSPGMVYIAPDSPLSIDPAIVS	IAVAGGLLIIFIVAVLIAYKRS	RESDLTLKRLQMQM	1278		
DB	1209	MVFPSPGSVS	VISLTLPAIISIAAGSGS	LLIIVILVLIAYKRSRENDLTLKRLQMQM	1268		
QY	1279	DNLESVALECKEAPAELOTHET	SDLDGAGIPFLDYRTVTMRVL	PPGIEDHPVLRLD	1338		
DB	1269	DNLESVALECKEAPAELOTD	NELTSDLLDRSGIPDYLDRIYAMRVL	PPGIEDHPVLREL	1328		
QY	1339	EYVPGYRQERVEGHLKFLAQL	INNKKVFLLSFIRTL	ESQSFMRDRGNVASLIMTVLOS	KL 1398		
DB	1329	EVQNGGQHV	EKALKFLAQL	INNKKVFLLT	FIRTL	ELQSFMRDRGNVASLIMTVLGQRL	1388
QY	1399	EYATDVLKOLLADLTDKNLES	KNHPKLLLRTE	ESVAERKMLTNWFT	FLLYKFLKECAGEPL	1458	
DB	1389	EYATDVLKQLLSDLTDKNLEN	KNHPKLLLRTE	ESVAERKMLTNWFA	LLHKFLKECAGEPL	1448	
QY	1459	FSLFCATKQOMKEGPDAIT	TEARYSIS	SEDKLIRQO	IDYKTLVLSVC	SPDNANSPEYVVK	1518

Db 297 GVEYRLIADAYLSPKRLAKELGIGSEREDILTFVPSQGNRIKPKKESVLCFLTKKI 356  
 Qy 364 NDRIKERLQSCYRGEGTDLWLKVKDI1PCSSALLTIDNFCGLDNAPLGVSDMVRGIP 423  
 Db 357 KDKIKERISQCYRGDGKSLPWLNLKELGICINSPLQIDNFCQDPNQPGGVTIEGTP 416  
 Qy 424 VFTEDRDMTSVIAVYKXHSFAFVTKSGKLKIRVDGPRGNA--LOVETVQVVDGPG 480  
 Db 417 LFLDKEDGMTSVAAYDYRGHTVVFAGTGRGVKILVDLSASSHLVQOVENVVHEGNA 476  
 Qy 481 VLRLMAFSKDHQOLYIMSRQLTRVPVSCGQYQSCGECGLSGDPCGCGWCVLHNTCKRE 540  
 Db 477 ILRLDLVSPDQYIYAMTEKQVTRVPVSCGQYQSCGECGLSGDPCGCGWCVLHNTCKRE 536  
 Qy 541 RCERSKEPRFAEMKOCVRLTVHPNNISVSOYNVLLVLETYNVPELSAGVNCETEDLSE 600  
 Db 537 KCRADLHRTSDQRCQVLTQVHPKNI SVTVSEVPVQAMVNDLSAGVNCSEFDFTE 596  
 Qy 601 MDGLVVGNOIQCYSPAACEVPRIIITENGHHVVLQOLKSKETGWTFASTSFVYNCVSHN 660  
 Db 597 MEGRIIDGKIYCTSPSAKEVPIPIRTHGDKRVVVKLYLKSKEGKFAVDVFYNCVSHQ 656  
 Qy 661 SCLSCVESYRCHWKYRVCHTDPKTSFQGRVKLPEDCPQLLRVDKILVPEVVKPI 720  
 Db 657 SCLSCVNGSFPCHWKYRVCHTDPKTSFQGRVNMSEDCPQLPSSQIYIPGVVKPI 716  
 Qy 721 TLKAKNLPQOSGORGECILANTQSGEORVPALRFNSSVQOQNTSYSEGMNINLPVE 780  
 Db 717 TLAKNLPQOSGORGECILANTQSGEORVPALRFNSSVQOQNTSYSEGMNINLPVE 776  
 Qy 781 LTVWNGHFNIDNPAQKVHLYKCGAMRSCGCLKADPDFACGCGQPOCQTLRQCPA 840  
 Db 777 LSVVWNGHFNIDNPAQKVHLYKCGAMRSCGCLKADPDFACGCGQPOCQTLRQCPA 836  
 Qy 841 QBSQWLELSGAKSKCNTPRITEIIPVTPREGTKVTIRGENLGLFREDTASHVKVAGVE 900  
 Db 837 LENPMWHAASRCDTPKTKLFPETGPRQGGTTLITGENLGLFREDTASHVKVAGVE 896  
 Qy 901 CSPLVLDGYIPAEQIVCEMGEA-KPSOHAGFVEICVACRPEFWARSOLYFMTLSD 959  
 Db 897 CVPVESEYISAEQIVCEMGEA-KPSOHAGFVEICVACRPEFWARSOLYFMTLSD 956  
 Qy 960 KPSRGPMSCGTOTVITGTLNAGSNVVMFGKOPCLFPHRRSPSVICNTTSSDEVLEMKV 1019  
 Db 957 TFSRGLSGTWISIEGNVNLNAGSDVSAIGRPFMSWRTAKEIRCKTPQGSTGKAEI 1016  
 Qy 1020 SVQVDRAKIHQDLV-FOYVEDPTIVRIEPEWSIVSGNTPIAVWTHLDLQNPQIRAKHG 1078  
 Db 1017 QILINRATWNSVHVNTEPTVQKIEPEWSIASGTPPLIVTGMNLATIKPKIRAKYG 1076  
 Qy 1079 GKEHINI CEVLNATEMTCQAPAL--ALGPDHOSDLTERPEERGFILDNVQSLILINKTN 1135  
 Db 1077 DVEKENNCTLYNDTWTWVCLAPSVDNPLRSPENGD---RPEIGFTMDNVHALLINTTS 1133  
 Qy 1136 FTYYPNPVFEAFPGSGILELKPCTPIILKGNLIPVAGNKNVTLVGEKPCPTVTVS 1195  
 Db 1134 FLYPPDPVEPLTAGNLELKPSPPLIKGNLIPVAGNKNVTLVGEKPCPTVTVS 1192  
 Qy 1196 DVQLLCEPNLIGRHKVMARVGMEYSPGMVYIAPDSPLSPLPAIVSIAGVGLLIIFIVA 1255  
 Db 1193 ETQLLCEPNLIGRHKVMARVGMEYSPGMVYIAPDSPLSPLPAIVSIAGVGLLIIFIVA 1252  
 Qy 1256 VLIAYKRSRESDLTKRLQOMQNDLESVALCEKFAELQDTHIELTSDLDGAGIPFL 1315  
 Db 1253 VLIAYKRSRESDLTKRLQOMQNDLESVALCEKFAELQDTHIELTSDLDGAGIPFL 1312  
 Qy 1316 DRYTYMRLVFPQIEDHPDLRDLVPGYRQERYEKGKLFQALINKNKVFLLSFRTLESQ 1375  
 Db 1313 ERYTYMRLVFPQIEDHPDLRDLVPGYRQERYEKGKLFQALINKNKVFLLSFRTLESQ 1368  
 Qy 1376 RPSFMRDRGNVASLIMTVLQSKLEYATDVLKQLLADLIDKNLESKNHPKLLLRRTESVAE 1435

Db 1369 RPSFMRDRGNVASLIMTVLQSKLEYATDVLKQLLADLIDKNLESKNHPKLLLRRTESVAE 1428  
 Qy 1436 KMLTNWTFELLYKFLKCECAGEPLFLFCALQOQKEGPIDAITGEARYSLSEDKLIROOI 1495  
 Db 1429 KMLTNWTFELLYKFLKCECAGEPLFLFCALQOQKEGPIDAITGEARYSLSEDKLIROOI 1488  
 Qy 1496 DYKTL-----VLSVSPDNANSPEVPVKILNCDTITQVKEKILDAI 1536  
 Db 1489 DYKTLNCPADDVGLSDESCCRSPQTLNCPNENAPPEIPVKILNCDTITQVKEKILDAI 1548  
 Qy 1537 PKVPCSHRPAKADMDLEWQSGGARMILQOEDITTKIENDWKRLNLTLAHYQVDPGSSVA 1596  
 Db 1549 YKGVPSYQRPKAGMDLEWQSGGARMILQOEDITTKIENDWKRLNLTLAHYQVDPGSSVA 1608  
 Qy 1597 LVSKQVYANVANNSTVSRTSASKYENMIRYTGSPDSLSRSTPMITTPDLESVQKMHVVK 1656  
 Db 1609 LVKQNSAYNINSGSTFTK-SLSRYESMLRTASSPDSLSRSTPMITTPDLESVQKMHVVK 1667  
 Qy 1657 NHEHGDQKEDGRGSKMSEIYLTRLLATKGTQKFXVDDLFETIFSTAHRSALPLAIKYM 1716  
 Db 1668 NHDLDQREGDRGSKMSEIYLTRLLATKGTQKFXVDDLFETIFSTAHRSALPLAIKYM 1727  
 Qy 1717 FDFLDEQADKHGHDHVRHTWKSNCPLPLRFWVMKIKNPQVFDIHKNSITDACLSSVAQ 1776  
 Db 1728 FDFLDEQADKHGHDHVRHTWKSNCPLPLRFWVMKIKNPQVFDIHKNSITDACLSSVAQ 1787  
 Qy 1777 TFMDCSTSEHRLGKSPSKNLLYAKDIPSYKNVVERYSYDIPKMPAISDQDMNAYLAQ 1836  
 Db 1788 TFMDCSTSEHRLGKSPSKNLLYAKDIPSYKNVVERYSYDIPKMPAISDQDMNAYLAQ 1847  
 Qy 1837 SRMHNEFTMSALSEIFSYVGYSEILGDLDDQCGKQKLAYKLEQVITLMSLDS 1894  
 Db 1848 SRLHLSQFNSMSALHEIYSYITKYRDEILTALEKDEQARRQRLSKLEQVIDTMAQSS 1905

RESULT 3  
 JC4980  
 plexin 1 precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 05-Nov-1999  
 R:Kameyama, T.; Murakami, Y.; Suto, F.; Kawakami, A.; Takagi, S.; Hirata, T.; Fujisawa, F.  
 Biochem. Biophys. Res. Commun. 226, 524-529, 1996  
 A:Title: Identification of a neuronal cell surface molecule, plexin, in mice.  
 A:Reference number: JC4980; MUID:96400291; PMID:8806667  
 A:Accession: JC4980  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-1894 <KAM>  
 A:Cross-references: DDBJ:D86948; NID:g1665756; PIDN:BAA13188.1; PID:d1013877; PID:g1665756  
 A:Experimental source: brain  
 C:Comment: This protein is a membrane protein, and plays a role in neuronal cell contact, sense of calcium ions.  
 C:Keywords: duplication; transmembrane protein  
 F:1-22/Domain: signal sequence #status predicted <SIG>  
 F:513-561,659-704,807-859/Region: cysteine-rich  
 F:1238-1264/Domain: transmembrane #status predicted <TM>  
 F:1266-1268/Region: hydrophilic

Query Match 63.4%; Score 6337; DB 2; Length 1894;  
 Best Local Similarity 63.5%; Pred. No. 0;  
 Matches 1209; Conservative 267; Mismatches 408; Indels 20; Gaps 13;

Qy 1 KAMPWNWTCLLSHLMVGMSSLTLLTRQAPLQKQBSFVTFRGEPAEGFNHLVVDERT 60  
 Db 1 MCLPPLSRLLLLLLLLLGRVWTAISPPAGLG-POPAPFTFVASD-WGTHLVVHEQT 58  
 Qy 61 GHIYLGAVNRIYKLSDLKVLVTHETGPDENPKYCPRIQVTCNEPLTTNNVNNKLLI 120  
 Db 59 GEVTVGAVNRIYKLSGNLTLLRAVTFGEVDNEKYPSPVQSCPHGLGSDTNVNNKLLI 118  
 Qy 121 DYKENRLIACGSYQGIKLLRLEDLFLKGLPYHKKHLYLSGVNESGVFGVIVS--YSN 178

119	DB	DY	ANRLLACGASQIQFRLDDLFKLBEPHRRKEHYLSVRREAGSMAGVLIAGPQG	178
179	QY	LQ	KLFIATAVDGKPEYFTTISRKLTKNSEADGMFPAYVFHDEFVASMIKIPSDTFTTIP	238
179	DB	QK	KLFIATAVDGKPEYFTTISRKLTKNSEADGMFPAYVFHDEFVASMIKIPSDTFTTIP	238
239	QY	DF	IYVYGFSNGFVYFLTLQ--PEWVSPGSGTTKEQVYVTSKLVRLCKEDAFNYSYEV	296
239	DB	AF	IYVYVFRSQFYLYTLQDLOLTSP--DAAGEHFTTSKIVRLCVNDPKFYSYVEF	296
297	QY	PI	CERSGVEYRLLQAAYLKSAGAVLGRITLVHPDDLLFTVFSKGQKMKSLDESALC	356
297	DB	PI	CEQAGVEYRLVDAYLSRPOALAKQGLAEDEEVLTVFAQOKXRVKPPKESALC	356
357	QY	IF	LKQINDRIKRLQSCYRGETLDLAWKVQDIFPCSSALLTIDNFCGLDMNAPLGV	416
357	DB	LFT	LAIRAKERIKERIOSCYRGEGLSLPMLLNKELGCINSPLQIDDDFCQDQFNQPLGG	416
417	QY	DM	VRGIPVFTEDDRMTSVIAYVKNHSLAFVCTGSKGLKIRVD--GPRGN-ALQYETV	473
417	DB	VT	IEGTPLVFDKEDGLTAVANDYQORTVVFAGTRSGRIRKILVLDANSGRPALAYESV	476
474	QY	QV	VDGPEVLDMAFSKDHEOLYITMSRQLTRVPVESCQYQSGCEGLGSDPHGCMVLH	533
477	DB	VA	QGNPILRLVLSPNRQVLYAMTEKQVTVPVESCQVYTSCELCGLSRDPHCGMVLH	536
534	QY	NT	CRKERCERSKEPRPASEMKOCVRLTVHPNNISVQVNVLLVLETYNVPELSAGVNC	593
537	DB	SI	CSRQDABEREPQFASDLQCVOLTQVPRNVSVTMSQVPLVLQANVPLDSAGVNC	596
594	QY	TF	DELSEMDGLVGNQIQCYSPAKEYPRIITENGDDHVVQLKSKETGMTASTFVF	653
597	DB	SP	EDFTETSEILEDRHCHSPSAREVAPITQCGQDQRVVKLYLXSKETGKFPASVDVF	656
654	QY	YNG	SVHNSCLSVESPYRCHWKYRVCTHDPKTCSEFQGRVKLPEDCCOLLRVDXILVP	713
657	DB	YNG	SVHOSCLACVNGSPFCHWKYRVCTYNNADCAFLGRVNMSEDCPOILPSTHIYVP	716
714	QY	VE	IKPTTLKAKNLPQSGORGVECTINTQSGEQRVPALRFNSSSVQCONTSYEGNE	773
717	DB	VG	VKPTTLARNLPQSGORGVECLFHPGPSARVTLRFNSSSLQCONSSYEGND	776
774	QY	INN	PVELTVVNGHFINDNPAQNVHLYKCGAWRSEGLCLKADPFCAGWCQGPQCT	833
777	DB	VSD	LPVNLVVVNGNFVIDNPQITQAHLYKCPALROSGLCLKADPFCGWCVAERRS	836
834	QY	LR	CHCPAE--SQWLELSGAKSKCTNPRTIIPVTGREGTKVTIRGENLGEFEDIAS	892
837	DB	LR	HCPADSPASWHAHSGSRCTDPKILKLSFETGPRQGTRLITGENLGLRFEDVRL	896
893	QY	HVK	VAGVECSPLVDGYITPAEQIVCEMGEAKPSQ--HAGFVEICVAVCRPEFMASSQIYF	951
897	DB	GV	HVKVLCSPVESEIYSAEQIVCEIGDASTLRAHDALVEVCVRDCSLHYRALSPPKFTF	956
952	QY	MT	LTLSDLKPSRGMSGGTQVTITGTLNAGSNVVMFGQPCLFHRRSPSYVCMNTSS	1011
957	DB	VT	PIFYRPSRPSLGGTIGIEGSHLNAGSDVAVISIGRCPSCFSWRNSREIRCLTPPG	1016
1012	QY	DEV	LEKMSVQVODRAKI--HODLVPOYVEDPTIVRIEPEWSIVSGNTPIAVMGTHLDLQ	1070
1017	DB	HT	PSAPIVININRAQLSNPEVKNYTIEDPTILRIDPEWSINSGGTLITVTGNLATVRE	1076
1071	QY	PO	IRAKGGKEHINI CEVLNATEMTQAPALALQPDHQSDLTERPEFPGILDNVQSLLI	1130
1077	DB	PR	IRAKYGGIERENSCWYNDTVMVCRAPSIDNPKRSPPELGERPDPDIFIMDNVFTLV	1136
1131	QY	LK	NFTYVNPVPEAFGPGSILELKPCTPIILKGNKLIPVAGNVKLANVTVLGEKPC	1190
1137	DB	LN	SFLYIPDVLPLEPSLGLLELKSSPILKGNLLPP-APGNSLNVTYVLGSTPC	1195
1191	QY	TV	TSVDQLLCESPNLIGRHKMARVAGMEYSPGKVIAPDPSLSPALVSIYAVAGGLI	1250
1196	DB	IL	TVSETOLLEAPNLITGCHVTVRAGFEFSPGMQVYSDSLTLPALVIGIGGGGLL	1255

## RESULT 4

JC4976

plexin 3 precursor - mouse

plexin 3 precursor - mouse  
C-species: Mus musculus (house mouse)

C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1996 #sequence revision 31-Dec-1996 #text change 05-Nov-1999

C;Date: 31-Dec-1996

C;Accession: JC4976

R; Kameyama, T.; Murakami, Y.; Suto, F.; Kawakami,

Biochem. Biophys. Res. Commun. 226, 396-402, 1996

A>Title: Identification of plexin family molecules in m

A:Reference number:

A;Reference Number: JC4976  
A;Accession: JC4976

A;Accession: JC4976

A;Status: nucleic acid s

A;Molecule type: mRNA

A;Residues: 1-1872 <KAM>

A;Cross-references: DDBJ:D86950; NID:g1655433; PIDN:BA...

C/Comment: This protein is a membrane protein with cell a

**C/Comment:** this protein is a membrane protein  
**C:Keywords:** duplication; transmembrane protein

C;keyworus: duplication, transmembrane protein  
E:1-19/Domain: signal sequence #status predicted <SIG>

```
F:1-19/Domain: signal sequence #status predicted
E:1216-1243/Domain: transmembrane #status predicted
```

F;1216-1243/Domain: transmembrane #status predicted <INM>

Length 1872.

Query Match	Score 59.9%;	DB 2;	Length 1872;
	59.9%;	DB 2;	Length 1872;

QY	67	AVNRIYKLSDDLKVLVTHETGPDNDPKCYPPRIVQTCNEPLTTNTNNNMKLLIDYKENR	126
Db	50	AVNRVFKLAENLTFLRAHVTPGIEDNARCYPYPPSMRVCSHRLVFDVNNVNNKLLIDYAARR	109
QY	127	LIACGSLYQGI CKLLRLLEDLFLKGEPPVHKKEHYLGSVNESGFGVYVYSNLDLKLFLA	186
Db	110	LVAAGSIWQIGI CQFLRLDDLLFLKGEPPHRRKEHYLSGAQEPDSWAGVILVEQVQFSPKLFVG	169
QY	187	TAVDGKPEYPTTISRRLTKNSBADGMFAFYVHDFEFVASKIPSDTFTTIIIPDFDIYVYV	246
Db	170	TAVDGKSEYPTTISRRLIDDEDSGMFLSVYQDEFVSSQIKIPSDTSLYPAFDIYIY	229
QY	247	GFSGPFVYFLTITQEPNVMVSPGSGTTKEQVYTSKLVRICKEDTAFNSVYVVPICERSGVE	306
Db	230	GFSVASFYFLTITQDITQOFTLLDTAGBKFTSKIVRMCAQDSSEFYSVYEPFICSGMRGVE	289
QY	307	YRLLOAAYLSKAGAVLGRITLVGHPDDLLFTVFSKGQKRKMSLDESALCICIFILKOINDR	366
Db	290	YRLVQSHLAKPGLLLAAQLGVPADEVDLFTIPSGQKQNRANPRQOTILCLFTLLSSINAH	349
QY	367	IKERLQCYRGEGTLDLAWLKVKDIIPCSSALLTIDNFCGLDMNAPLGVSDMVRGIPVFT	426
Db	350	IRRIQCYRGEGTALPWLNLKELPCINTPLQINGNFCGLVLNQPLGLLHVIEGLPLLA	409
QY	427	EDDRMTSVIAYVYKXHSLA FVGTGSKGLKKIRVDGPRGNALQYETVQVVDPGPVLDMA	486
Db	410	DSTDGMASVAAAYTYHQHSVVYFIGTRSNLKKVRVDGSQ-DAQUYETVSVVQSGPILRDL	468
QY	487	FSDHEOLYIMSBROQLTRVPVESCQYOSCGECLGSDPHCGMVCVLHNTCTRXKRCERSK	546
Db	469	FSPDHRHIYLLSEKQVSQLPVEICEQYLSACALGSGDPHCGMVCVLOHRCRGACFGAS	528
QY	547	EPRRFASMKQCVRLTVHPNNISVQSYNNVLLVLETYNVPELSAGVACTPBDLSEMDGLVV	606
Db	529	APHGFAELSKCIQVRVRPNNVSVTSGVQLTVAMRNVDPDLSVGVSCSPREVTESEAIL	588
QY	607	GNQIOCYSPAAKEVPRIITENGDDHVVLQLSKETGMTFASTSVFYVNCVHNSCLSC	665
Db	589	PSGELKCPSPLOELQTLTRGHGATHVRLQLLSMETGVRFAGVDFVFNCSALQSCMSC	648
QY	666	VESPYRCHWKYRHVCTHDPKTSFQBEGRVKLPEDCPQLLRVDKILVPVBEIKPITLAK	725
Db	649	VGSPYPCWKYRHVCTSHPECSFQBEGRVHSPEGCEIILPQGDLLIIPVGMQPLTLRAK	708
QY	726	NLPQOSGQRYGECILNIQSGEORVPALRPNSSVQCONTSYSEGMENINLPELTVVW	785
Db	709	NLPQOSGQKXNYEVRVQGRHVRPAVRPNSSSVQCNASIFYEGDEFDGTDLDFSVVW	768
QY	786	NGHENIDNPAQNVHLKYCGAMESCLCLKAPDDFACGQCQPGQCTLRHCPAQESOW	845
Db	769	DGDFPIDKPSFRALLYKCAWRQPSGCLCUKADPRFNCWCI SEHRQLRAHCPAPKSNW	828
QY	846	LELSGAKSKCTNPRITBIIIVTPGREGTKVTIRGENLGLIEFRDIASHVAVAGVECSPLV	905
Db	829	MHPSQKGARCSHPRTIQHPLTGPKEGTRTVIVGENLGLTSREVG--LAVAGVRCNSIP	886
QY	906	DGYIPABQIVCEMGEA-KPSQHAGFVEICVAVCRPEPMARSSQLYPMTLTLSDLKPSRG	964
Db	887	TEXTVSABRIVCEMBEESIVPPPGPABELCVGDCSDAPRTQSQOILYSFVTFTFDRVSPSRG	946
QY	965	PMSGGTQVTTTGTNLAGSNVVMFGQPCLEHRRSPSYIVC-NTWSSDEVLEMKYSVQV	1023
Db	947	PASGTRKLTISGLSDAGSRVTVIIRDGECQFVRRDAEAVICLISPVSTLGFQSPITLAI	1006
QY	1024	DRAKI-HODLVFOYVEDPTTIRIEPEWSIVSGNTPIAVMGTHLIDLIONPOIRAKHGKHEH	1082
Db	1007	DHANISNTGVITYTQDPTVTHLEPTWSIINGSTSIIVSGTHLLTVOEPRVAKRYGIET	1066
QY	1083	INICEVLNATEMTCQAPALAGPDHQSDLTERBEEFGFILDNTVQSLIILNKNTFTYYPNP	1142
Db	1067	TNTQOVINDTAMLCAPGIEFLGHQPRAQOGEHPDEFGLLDHVQAARSLNRSSFTYYPDP	1126

Qy	1143	VFEAFGSGILELEKPGTPIILKGNLIPPVAGGNVKNYITVLVGERPCTVTVSDVQLLCE	1202
Db	1127	SFEPFGSGVLDVKPQSHVVLKGNLIPPAAGSS-RLNYITVLVGGOPCALTVSDTQLLDC	1185
Qy	1203	SPNLIGHKVMARVGMWYSPGMVYITAPDSPLSPATVSTAVAGGLIIFIIVAVLTAAYKR	1262
Db	1186	SPQSTGHOPWVLVGGLEFWLGLTHITADRALITLPAWGLAAGGGLLLLAITVVLVAYKR	1245
Qy	1263	KSRESDLTLKRLQWMDNLESVALCEKFAELQTDIHELTSDLDPGAGIPFLDYRTYTM	1322
Db	1246	KTQDADRTLKRLQLOMDNLESVALCEKFAELQTDINELTNHMDGVQIPFLDYRTYAV	1305
Qy	1323	RVLFPGETHDHVLRLDLEVPYGRQERYEKGILKPAQLINNVKVFLLSFTIRLESQRSFSMRD	1382
Db	1306	RVLFPGETHAHPVLKELDTP----PNVEKALRFLGQLLHSAFLLTIRHTLEAQSFSMRD	1361
Qy	1383	RGNVASLITMTVLOSLEYATDVLKOLLADILDKNLESKNHPKLLLRRTESVAEKMLTNWF	1442
Db	1362	RGTVASLITWVALQSRDYATGLLKOLLADILEKNLESKNHPKLLLRRTESVAEKMLTNWF	1421
Qy	1443	TFLAYKFLCEKAGEPLFSFLCAIKQOMEKGPIDAITGEARYSLSEDKLIRQQIDYKTLVL	1502
Db	1422	TFLLHKFLCEKAGEPLFLLYCAIKQOMEKGPIDAITGEARYSLSEDKLIRQQIDYKTLTL	1481
Qy	1503	SCVSPDNWNSPEVPVKILNCOTITOVKEKILDALFKNVPCHRPKPAADMLEWRQSGAR	1562
Db	1482	HCVCPESSGAQVPVKVNLNCDSITQAKDKLLDVTYKGIPIYSORPKAEDMDLEWRQGRMAR	1541
Qy	1563	MILQDEDITTKIENDWKRLNTLAHYQVPGDSVVALYSKVQTYANAVNNSVTVSTASKEYE	1622
Db	1542	LILOQDEDITTKICDMKRVNSLAHYQVTDGSLVALVPQVSAVNMANSFTFTIR-SLSRYE	1600
Qy	1623	NMIRYTGSPDSLSRSTRPMITPDLESVGKMMHVLKVNHEHGDQKEGDRGSKMVSEIYLTRL	1682
Db	1601	SLLRPAASSPDSLSRAPMLTPDQEGAGTKLWHLVRNHDHTDHREGDRGSKMVSEIYLTRL	1660
Qy	1683	ATKGTLOKQFVDDLPEFETIPSTAHRGSALEPLAIKYNFOPFLDEQAOKHGHIHDPHVHTWKSNC	1742
Db	1661	ATKGTLOKQFVDDLPEFVFTSAHRGSALEPLAIKYNFOPFLDEQAQORQISDDPDVHTWKSNC	1720
Qy	1743	LPLRFWNMLKNQOPFVDIHKNSITDACLVSVAQTWMDSCSTSEHRLGKDSPSNKLLYAK	1802
Db	1721	LPLRFWNVLKNQOPFVDIHKNSITDACLVSVAQTWMDSCSTSEHRLGKDSPSNKLLYAK	1780
Qy	1803	DIPSYKNWERYTSDIGKMPAISDDQMNAYLABQSRMHNEFTMTSALSSEIFSVGVKEYE	1862
Db	1781	DIPNYSWERYRYADIAKMASISDDQMDAYLVQSRLLHANDFNVLVSALSSELYFVVTKYRQ	1840
Qy	1863	EILGPLDHDQCGKQKLAUYKLEQVITLMSLDS	1894
Db	1841	EILTSLDRDASCRAKHKLROKPEOITITLVSSSS	1872

## RESULT 5

TL13937  
 Plexin A - fruit fly (*Drosophila melanogaster*)  
 C:Species: *Drosophila melanogaster*  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
 C:Accession: TL13937  
 R:Winberg, M.L.; Noordermeer, J.N.; Tamagnone, L.; Comoglio, P.M.; Spriggs, M.K.; Tessie-Cell 95, 903-916, 1998  
 A:Title: Plexin A is a neuronal semaphorin receptor that controls axon guidance.  
 A:Reference number: Z17621; MUID:99091049; PMID:9875845  
 A:Accession: TL13937  
 A>Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1945 <WIN>  
 A:Cross-references: EMBL:AF106932; NID:g4056673; PID:g4056674; PIDN:AAD09425.1  
 C:Genetics:  
 A:Gene: plexA  
 A:Cross-references: FlyBase:FBgn0025741  
 A:Map position: 4  
 C:Function:

A,Description: may function as repellents during axon guidance  
C,Keywords: cell adhesion; nerve

Query Match	36.3%;	Score 3631;	DB 2;	Length 1945;
Best Local Similarity	41.9%;	Pred. No. 2.4e-226;		
Matches	828;	Conservative 320;	Mismatches 638;	Indels 190; Gaps 60;
QY	26	LTRQAPLQKQSFYTFRGEPA--EGFNHVVDERGTHIYLGAVNRIYKSLSDLKVLVT	83	
DB	48	LTNANAPIKNAKMLNITNVAADFYKLNHLLVDTITGRVFVGVGNRLYQLSDLELSET	107	
QY	84	HETGPEDEPNKCYPPRIVQTC--NEPLTTNNVNMKLLIDYKENRLIACSLYQGIKLL	141	
DB	108	VKTGPQNDSEVC-----SILDCPLNAVRSPTDYNKVLIDRATSRLLIACGSLFGCTCTVR	163	
QY	142	RLBDFKLGEPYKHKHYLSCVNESGVFVIVSYNSLDDKLFATVADGKPEY-----	195	
DB	164	NLQNSII-----EHEVPD-----AVANDANSSTVAFIA-----PGPPQHPVTNVM	205	
QY	196	-----PPTISSKLTKNSEADGMFAYVFHDFVASMIKIPSDTFT-----II	237	
DB	206	YGVVTYNNSPYSEIAPAVASRLKTK-----MF-----QIASSAVTTGTRFINSYAR	255	
QY	238	PFDIYVYVFGSGNFYFTLQPEMVSPGCTTKQVYTSKLVRLCKEDTAFNSYVEVP	297	
DB	256	ETVFNVYVFGSSERFSYFLTQ--LKSHSHSPKE--YITKLVRICQEDSNYSYZTEIP	311	
QY	298	IGC---ERSGVEYRLQAAVYLSKAGAVLGTGLVHPDDLLFTVFSKQKRMKSLDESA	354	
DB	312	VECISDAQGKTENLVQAGFLGRPSSNLAQSLGISQNDVLFAVSKGEGN--TPTNNSA	369	
QY	355	LCIPILKQINDRIKELQSCVRBGTLDLAWLKVDIPC--SSALLTIDDDPCGLDMAPL	413	
DB	370	LCIYSLKIRKFWQNIKSCFNGSGWGLAFIS--PSMPCVLTKLTQIGEDFCGLDVNSPL	428	
QY	414	GVSMDVRGIPVFTEDRDMTSVIAVYVKNHSLAFVGTGSKGLKIRVDPGRNALQVETV	473	
DB	429	GGETPITSVPV--AMFNKLTLSVAATSTSGYTVFVGTSDGFLKVKVIES--SSIANEYASF	486	
QY	474	QVDPDGVPLMDAPSKDEHQIYINSEQLTAVPYVESCGYQSCGECGLSGDPHCGWCVLH	533	
DB	487	AVDLGSEINDQPDQNLXIYVMSKTKVVKVDFCDSDYKTCGDCGLGARDPYCGWCSLE	546	
QY	534	NTCTRKERC--ERSKEPRRFAS--EMKQCVRLT--VHPNNISVSQYVNL--LVLETYNVPELSA	589	
DB	547	NKCSPRNCCDDANDPFVWSYKTKCTTITSVVPHQI-ORTTARTLELIID--HLPOLKE	604	
QY	590	GVNCTFEDLSEMDGLVVG-----NQICYSPAACEVPRITENGDDH--VVOQLKSKETG	643	
DB	605	NLICAF--TTEDKALFTNATKRRGVNCTTPTDMLPQI--EOGKHFTAKLSVRTR--NG	659	
QY	644	MTFASTSFVFNCSVHNSCLSCVSPYRCHCKYRHHVCTHD--PKTC-----	688	
DB	660	PDVYSTDTFFDCDTHSSCTCVSSEFPDCMVCZAHRCTHDTAENCRNDIILVTGVSRLGP	719	
QY	689	SFQGRVKLPDPCQLLRRV---DKILVEVEVIKPTTLKAKNLPQPSQGRGYECILNIQG	745	
DB	720	SYRSG-----PGFCTINATGDSSEVLVAGTSSKIKVKVHIIGQ--FIVQTRFVQCFNIEG	774	
QY	746	SEQRVPAL--RFNSSVQONTSYSYEGMEINNLPELTVVWNGHFNIDNPAQNKVHYLK	803	
DB	775	---RVTSLNAQLLGDITYICDSMEFYQTSRS--PLNTATPAVINGGSKPLDNPHNIHVYIR	830	
QY	804	CGAMRESGLCLKADPDFACGWCQPGOCTLRHCPAQ---ESOWLELSGAKSKCTNPRI	860	
DB	831	CREWADSGCICLASEKYNWCSCSTNTCEVEEQCNKKEGKTDWLNRS---EICPNPEI	887	
QY	861	TEIIPVTGPREGGTKVITIRGENGLGLEFRDIAHVKAVGECSPVLDGYIPAEQIVCEM--	918	
DB	888	HTFQPKTGPWEGGNTITIRGINLGKNYNDIYSGVRIAGINCMPFPQFVIDTKVICVTVDS	947	
QY	919	-GEAKPSQHAQFVEICVACRPEFPAWRSQLYFMTLILSLDKPSRGWMSGTQVITIGT	977	

RESULT 6  
T13164  
plexin B - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
C:Accession: T13164

DB	948	PGQMVRNGKIVVQI-----GDYRGESKEDYEFVDPKILDNFPKFGPTSGGTEIHITGK	1001	
QY	978	NLNAGSNVVMVFGKQ--PCLFHRSPSYIVNCTTSSDEVLEMKVSVQVDR--KIHQDLVQ	1035	
DB	1002	HLNAGSRIQASINDHLPCKILSTDSSQAI CRTSAPGIIIEGLKMSFDNGPREFNDFYFK	1061	
QY	1036	YVBDPTIV-----RIBPEWSIVSGNTPIAVWGHLDLIONQIRAKHKGKHEINICE	1087	
DB	1062	YVLDPVTVEHVSSPGSQIKVPKGPAGIRISVTGTQFTSIQNPNIVYVYNGEMYASPCR	1121	
QY	1088	VLNATMTCOAPALAGDPHQSDLTERP--EEFGFILDN---VOSLLILNKTNTFYYPNP	1142	
DB	1122	VQSDTEMECASPVDV--DSHVIAERPIILEYGLMDNVLRVQNLSKVNNHFLYFNP	1179	
QY	1143	VFBAGFSGILELKPPTIILKGNLIPPVAGGNVKNLYTVLVGKEKPTVT--VSDVQLLC	1201	
DB	1180	EYFIFER--VKYFKSYLVLTNGRNLDRACKESDVE---VKIGNGFCNITSLRQQLTC	1233	
QY	1202	ESPN-----LIGHKVMARVG--GMEYSPGMV--YIAPDSPLSLPAIVSIYVAGGLLIIFI	1253	
DB	1234	RPPEATATKSMNGPEVIVRIGTSLRYIGLSYESSNIILDWGENVIFAVIATVILL	1293	
QY	1254	--VAVLTAAYKRESDELTLKRLQOMDNLSRVALECKEAFELQTDIHELTSLDGAG	1311	
DB	1294	IFVALLVAYKKSSERSVLRNMQEOMDILELRVAECKEAFELQTEMTDITGLTSGG	1353	
QY	1312	IPFLDYRTYTMVFLFPGIEDHPVLRDLEVPYGRQ--RVEKGLKFLAQLINNKVLLSFI	1369	
DB	1354	IPFLDYSYAMKILFPNHEDHIVLQ-----WERPELLRKEKGLRIFGQLINNKTEFLFI	1408	
QY	1370	RTLESQSFMRDRCNVASLIMTVLQSKLEVATDVILKOLLADLIDKNLSKNHKLRLR	1429	
DB	1409	RTLESNYFYSRERNVNASLIMTVLQSKLEYCTDILKTLGLDIEKIEGKSHPKLLRR	1468	
QY	1430	TESVAEKMLTNWFTFLYKFLKECAGEPLSFLCAIKQOMEKGPDAIDAITGEARYSLEDK	1489	
DB	1469	TESVAEKMLSNWFTFLYKFLKECAGEPLMFLRAVKGVQDKGPDVACTHEARYSLEBK	1528	
QY	1490	LIRQIDYKTL-----VLSCVSPD--NANSEPVVKILNCDTITQVKEKILDAIFK	1538	
DB	1529	LIRQSIDFRPMNVYASIIQPIFCNNLDMLPSHTENSVKVLDCDTIGQVKEKCLTIYR	1588	
QY	1539	NVPCSHRPAADMLEWQSGARMLODEDITTKIENDWKELNLAHVQVDPGVSVALV	1598	
DB	1589	NIPSPRPDKDLLEWMTGATGRVILYDEDTSTSEWKKLTLQHNVPDAGLSLV	1648	
QY	1599	SKQVTAVNAVNNSTVSRTSASKYE--NMIRYT--GSPDSLSRSTPMITPDLESVGMHLV	1655	
DB	1649	PQSSINFSILSDKNEKS--HKYETLNISKYTSSTPTFSRAGSPLNNDMHENGLRYHLV	1707	
QY	1656	KNHEHGDQKEDGRGSKMSEIYVLTLLATKGLQKFDVDDLFTIISTAHGSLPLAIKY	1715	
DB	1708	KHSDMDQKEGERNVKNLVSSEIYVLTLLATKGLQKFDVDDLFTIISTAHGSLPLAIKY	1767	
QY	1716	MFDELDEADKHGHDHPVHRTHTWKNCLPLRFVWNNIKNPQFVFDIHKNSITDACLSSVA	1775	
DB	1768	MFDELDDQAQHGHTDPEVHTWKNSLPLRFVWNNIKNPVFDIHKNSIVDSCLSVA	1827	
QY	1776	QTFMDCSTSEHRLGKSPSNKLLYAKDIPSYKNVYRYISYDYGKMPAISDQDMNAYLAE	1835	
DB	1828	QTFMDCSTSEHRLGKSPSNKLLYAKDIPSYKNVYRYIRDIRMSPISDQDMNAYLAE	1887	
QY	1836	QSRMHMNFNTMSALSEIFSYVGVKYSSEILGFLDHDQCKQKLYAKLQVITLMS	1891	
DB	1888	ESRLHTTEFTNICALHELYTAVKYNEQLTVTLEEDFESQKQRLAPKLBQVHIMS	1943	

[illegible]

1851 FSIILTV---NEELPAVKWLFDDLLDEARRHXIADTDIVHAWKSNCLPLRFWVFIKN 1907  
1755 PQVFDIHKNSITDACLSSVVAQTFMDSCTSEHRLKGDSPNKLKLYAKDIPSPYKQWVERY 1814  
1908 PDFIDVKNKTSVDSCLSVXAQTFMDACSTSEHRLKGDSPNKLKLYAKDIPSPYKQWVERY 1967  
1815 YSDIGKMPALSDOMNAYLAQSRMHNEFTWALSSEIFSYGKYSEETLGPLDHDHDOC 1874  
1968 YRDVSLPQISDQEMSTAMQOLSVRQNEEFTDTSALKELYIYITKYKQDMESLETINC 2027  
1875 GKOKLAVKLEQVITLMSLD 1893  
2028 RKMHLRKLGNVAATLGD 2046

RESULT 7  
T23298  
hypothetical protein K04B12.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T23298  
R:Steward, C.  
submitted to the EMBL Data Library, December 1996  
A:Reference number: Z19723  
A:Accession: T23298  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1806 <MIL>  
A:Cross-references: EMBL:Z83232; PIDN:CA805755.1; GSPDB:GN00020; CESP:K04B12.1  
A:Experimental source: clone K04B12  
C:Genetics:  
A:Gene: CESP:K04B12.1  
A:Map position: 2  
A:Introns: 24/2; 77/3; 294/2; 378/3; 433/3; 478/3; 523/3; 743/1; 794/3; 1049/3; 1414/3;

Query Match 16.1%; Score 1609.5; DB 2; Length 1806;  
Best Local Similarity 26.4%; Pred. No. 2.6e-95;  
Matches 525; Conservative 346; Mismatches 773; Indels 341; Gaps 81;

11 LLSHLMVGMSSILLTROPAPLQKQKRSVTFRCEPAEGNHLVVDERTGHVILGAVNR 70  
37 LISHFLRA-----VTQPPPETEGVKOKLPHFSGH-----IDDFIVSRDQOTIYVASLNR 85  
71 IYKLSDDLKVLVTHET--GPDEDNPKC-----YPPRIVQTCNEPLATTN 112  
86 LTSLSIS-NFSIQHEVSLGPQDSPWCSADGSKLKNQHHVFQIIIFRDNRPFT-- 142  
113 NVNKMLLIDYKENRLIACSLYQGIKLLRLLEDLPKLGEPYHKKHLYSGVNESGSVFGV 172  
143 DVRTKILQILPTNQLQCGSVKLGSC-----STFNSK-----LSLITES-----T 182  
173 IVSYSN-----LDDKLFIAVADGKPEY--FPTTISRKLK-KNSEADGMF---A 215  
183 IAVANSPDASTVKIIONRLIIVASATKESPYRPPFAVATRNLPGLNVENAGDLEGEA 242  
216 YVFHDEFVAMIKIPSDTFTIIPDFIYVVGSSGNFYFLTLOPEVMVSPGSTKQV 275  
243 AVFLRAAYKNAFK-----FLYTFTHQHFVFA-----MVTRESRLP--- 280  
276 YTSKLVRICKEDTANSVVEVPIGE-RSGVEYRILLOAYLSKAGAVLGRITLVGHPDDDL 334  
281 MTTRILIRECRNDTKFESYSEILOCRGENDNTNYPFLNALIQSY-----DK 325  
335 LFTVFSKQKRWKSLDESALCIFILOKINDRIKERLOSQCYRGEGLDLAWLKVKDIPCS 394  
326 LIASFST-----STSPKSSICVFSNQKVKLFTWNVNDCRSTDSIRLPHIG-RDTKC- 378  
395 SALLTIDNFCGLDNAPLGVSDMVRGIPVFTED-RDRMTSVIAYVYKXHSIAFVGTGSG 453  
379 KAHIPLDEDC-----ELGVGSGIELVEMSTKDMGVKTSIMAV-----DQKAIFAFTTS 429  
454 KLKKLRVQDPRGNAL-QYETVQVVD--PGPVLRLDMAFSKDHEQLYMTSBRQLTRVPVESC 510

430 QIVMFKWDEHHSNQLBEEYGRKEVGDGRTSGSEVSKWV--KFGDFVIVQMPYGIILIELSTC 487  
511 GOYQSGCEGLSGDPHGCGWCVLHNTCTRKERCERSKEPRPASEMKQC--VRLTVHPNNI 568  
488 SHSSCTCELVSDPLCQWCHPTQSCCTTSARCT-----SPVTSQCPIVDGDPISIV 539  
569 SVSQYNNVLLVLETYNVPELSAGVNCCTFE-----DLSEMDGLVGNQIQCYSPAACEVPRI 623  
540 SVNSSTPI---SFNIHILPPVPGFTYRCQFGTSTSSIKANWTTTGVSCPS-----EI 588  
624 ITENGHHVVLQQLKSKETGTFASTSFVFNCSVHNSCLSCSPYRCHCKYKHVCTH 683  
589 FTSPTNTEIILLTTSINN---PISRRHNTVYDCCGYGTCCSCMSSEYNCAMCSGLHKCSN 645  
684 DPKTCSFQEGRVKLPEDCPQLLRVDKILVPVEV--IKPITLKAKNLPOQSGQRGYE--C 739  
646 ---SC-----GALEKSKAC---VKIQPMRLPAIGSQEQEIVLEASN---DTLDRRHEHFC 692  
740 ILNIQSGEQRVPALRFNSSVQOQNTSYSYEGMEINNLPLVELTVVW--NGHFNIDNPAQNK 798  
693 KVN-----EQVSLAKIASDSIRCCKIQLTSLNNTSANMVVPLSLITRDSVIDIAN----- 742  
799 VHLKCGAMRESGCLCLKADPDFACGCGQPGQCTLRQCPAESQWLELSGAKSKCTNP 858  
743 VLSYCTNLASDCSSCLALSPLSCGWCN--RQCSHECH-----ESK-----ATAVCDPP 790  
859 RITEIIPVTGREGTKVTIRGENLGLFEDRIASHVKVAGVECSPLVDGYPABQIVCEM 918  
791 RIDKFEPTSGPIEGGTIIKIYGNLGMVSDVRGKIYVAGSRCN--IVEYHYVSNMIACQV 848  
919 GEAKPSQAGVETICVACRPEFPMARSQLYFWMTLTLSDLKPSRPMGSGGTQVITIGN 978  
849 DKGYSV---GPISRVSGRATVA--VAESSELSYFVSTISFSAPLYGPIGSGGRITLYGQN 904  
979 LNAGSNVVMFGKOPCLFHR-RSPSYIVCNTTSSDEV-LEMKVSVQVDRAKIHQDLVFOY 1036  
905 LSSQSQTSVTYVGMPCPIERNVNSVTLCITPSGTRIGKARVAVVVDHSQTLQDPFEY 964  
1037 VEDPTIVRIEPEWSIVSGNTPIAVWGTDLHLQIONQIRAKHGKKEHINI-----CEVLN 1090  
965 RSDPSISIFPMTSPKAGGRIVYVQGNLNTVQTAKFLISSPTPPFYIISLAPCHIN 1024  
1091 ATEMTQAPALAGDPDHQSDLTERPEEFGFILDNVQSLILNKNTFTYYPNPVFEAFGPS 1150  
1025 STLMTQMTPK-----LLETITRVEYTRQPMGIYPNPLSPF--K 1062  
1151 GILELKGPTPIILKGNLIPPVAGGNVKL---NYTVLVGKEKCTVTVSQV-QLLCSPNL 1206  
1063 GVRYHQGEQSLILEGHNL-----NLAAEPNDKIFIGNERCYVTLVDVRQVJCSGP-- 1113  
1207 IGRHK-----VWARGGMEYSQWVYIAPDSPLSLPAIVSIAGVGLLI 1251  
1114 VRQPKATDERGIPINGDNPPLVTIVIGSLRMEGLIEY---SHALPSRLSLLILG--LLL 1168  
1252 FIVAVL-----IAYKKSRESDLTLKRLQMDNLESRLVALECKEAFELQTD--IHELTS 1305  
1169 FIVVTLVWCLVFKRRQREKEKYRKIQLOMENLENNVRKECKQAFELQTNLVLSPKSA 1228  
1306 DLDGAGIPFLDYRTYTMVLFPP--GIEDHPVL-RDLEVPGYROERVEKGLKLAQLINNK 1362  
1229 NSVNLGPELINPFPFVFNLLWSDNNLTSAPLARTLFT-----LAQPHALLSFK 1278  
1363 VFLLSFIRLTESQSFMSMEDRGNVASLIMTWLQSKLEYATDVLKQLLADLIDKNLESKNH 1422  
1279 GIFTIVEAASDVSLSTSEKSLASLLISVLLRNFSYCTEVVVDLLRAHARSQNK-R 1337  
1423 PKLLLRRTSEVAEKMLTNWFTFLLYKFLKECAGEPLFSFCAIKQOMEKGPIDAITGEAR 1482  
1338 AELLFNSDSVVVEKMSKWSICLYSHLTQPMNS-YFYLYKALQYQTDKGFVDAVTGDAR 1396  
1483 YLSLEDKLIROQIDYKTLVLSVSPNPNASPEVPVKILNCDDTITOVKEKILDAIFKNVPC 1542  
1397 YTINEAKLRRESVDTKLIR-VIPFEKDESIDLEHVACDAICQVKQKVASAVYRETPY 1455



1339 EYPGYRQERVEKGLKFLPA--QLINNVKVFLLSFRTLESQRSFMRDRGNVASLIMTVLOS 1396  
1042 ----HNRDANDKNESUTALDALLCNKSFVTVHTLEKQNFVKORCLFASFLTALQT 1097  
1397 KLEYATDVLKQLADLIDKNLESKNHPKLLRRTESVAEKLTWNFTFLYKFLKECAGE 1456  
1098 KLVYLTSLILEVTRDLMEQC--SNMQPKMLRRTESVVEKLLTNMWSVCLSGFLRETVE 1155  
1457 PLFSLFCAIKQOEKGPDAITGEARYSLSEDKLIQQIDYKTLVLSCV---SPDNANSP 1513  
1156 PFYLLVTLNKNKNGPVDVITCKALTYTLNEDMLLMQVPEFSTVALNVVFEKIPENESAD 1215  
1514 ---EVPVKILNCDTITQVKEKILDAIPKVPNCSPHRPKAADMDLEWRQSGARMILODEDI 1570  
1216 VCRNISVNVLDCTIGQAKKEKIFQAFSLKNGSPYGLQNLGIEGLQWTRQKELLDDISS 1275  
1571 TTKIENDKRLANTLAHVQVDPGVSVALVSKQVTAYNAVNNSTVSRTSASKYENMIRVTGS 1630  
1276 SVILEDGITKNTIGHYEISNGSTIKV-----FKKIANFTSD 1312  
1631 PDSLRSETPMITPDLESQVKNMHLVKNHEHQDQKEGDRGSK--MVSEIYLTRILATKGTIL 1688  
1313 VEYSDHCHILIPDSEA-----FQDVQGRH-----RGKHKPKVKEMYLTKLLSTKVAI 1361  
1689 QKFVDDLFEITFSTAHRGSSALPLAIKTMFDLDRQADKHGIDHPVHRTWKSNCPLRPF 1748  
1362 HSLVKLFRIWSLPLN--SRAPFAIKYFFDFDLQAQENKKITDPDVVHIWKTNSLPLRF 1419  
1749 VNMKNPQVFDIHKNITDACLSSVAGTQFMDSCSTSEHRLGKDSPSNKLLYAKDIPSYK 1808  
1420 VNLKNPQVFDIKTKPHIDGCLSVIAQAFMDAFSLTEQQLGKAPTNNKLLYAKNDIPTYK 1479  
1809 NWVERYSYDICKGMPAISQDQDMNAYLAQSRMHMEFTMSALSIFSVGVKYSBEILGPL 1868  
1480 EEVKSYYKAIKRDLPPLSSSEMEFLETKESKHEFEFEVALTIIYKIVYKYPDEILNKL 1539  
1869 DHD---DQCGKQKLAYKL 1883  
1540 ERERGLEEAQQLLHVK 1557  
RESULT 9  
T46426  
hypothetical protein DKFZp434G0625.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C:Accession: T46426  
R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23028  
A:Accession: T46426  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-317 <AAA>  
A:Cross-references: EMBL:AL137352  
A:Experimental source: adult testis; clone DKFZp434G0625  
C:Genetics:  
A:Note: DKFZp434G0625.1.  
Query Match 13.2%; Score 1322; DB 2; Length 317;  
Best Local Similarity 97.0%; Pred. No. 5.6e-78;  
Matches 256; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
206 KNSADGMFAVYFHDHFVASMIKIPSDTFTIIPDFDIYVYGFSSGNFYFLTLPQEMVS 265  
1 KNSADGMFAVYFHDHFVASMIKIPSDTFTIIPDFDIYVYGFSSGNFYFLTLPQEMVS 60  
266 PPGSTTKSQVYTSKLVRLCKEDTAFNSYVEVPDIGERSGVFRLQAAVLSKAGAVLGT 325  
61 PPGSTTKSQVYTSKLVRLCKEDTAFNSYVEVPDIGERSGVFRLQAAVLSKAGAVLGT 120  
326 LGVHPDDDLLFTVFSKGQRKMKSLDESALCIFIILKQINDRIKRLQSCYRGEGLDLAW 385

121 LGVHPDDDLLFTVFSKGQRKMKSLDESALCIFIILKQINDRIKRLQSCYRGEGLDLAW 180  
386 LKVKDIPCCSALLTIDNFCGLDMNAPLGVSDMVRGIPVFTEDRMTSVIAYVYKNHSL 445  
181 LKVKDIPCCSALLTIDNFCGLDMNAPLGVSDMVRGIPVFTEDRMTSVIAYVYKNHSL 240  
446 AFVGTSGKGLKKTIVDQPRGNALQ 469  
241 AFVGTSGKGLKKSFGTPOGGITQ 264  
RESULT 10  
T47133  
hypothetical protein DKFZp761P19121.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47133  
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24374  
A:Accession: T47133  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-177 <AAA>  
A:Cross-references: EMBL:AL162013  
A:Experimental source: adult amygdala; clone DKFZp761P19121  
C:Genetics:  
A:Note: DKFZp761P19121.1  
Query Match 7.7%; Score 766; DB 2; Length 177;  
Best Local Similarity 79.1%; Pred. No. 2.3e-42;  
Matches 140; Conservative 20; Mismatches 17; Indels 0; Gaps 0;  
1718 DFLDEQADKHGIDHPVHRTWKSNCPLRPFVWVMIKNPFQVFDIHKNITDACLSSVAGT 1777  
1 DFLDEQADKHGIDHPVHRTWKSNCPLRPFVWVMIKNPFQVFDIHKNITDACLSSVAGT 60  
1778 FMDSCSTSEHRLGKDSPSNKLLYAKDIPSYKNWVERYSDICKMPAISDQDMNAYLAQSR 1837  
61 FMDSCSTSEHRLGKDSPSNKLLYAKDIPSYKNWVERYSDICKMPAISDQDMNAYLAQSR 120  
1838 RMHMEFTMSALSIFSVGVKYSBEILGPLDHPDQCGKQKLAYKLEQVITLMSLDS 1894  
121 RLHLQFNSMSALHIIYSYITKYDEILAALEKDEQARRQLRSKLEQVVDVTMALSS 177  
RESULT 11  
JC5148  
hepatocyte growth factor receptor precursor - African clawed frog  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) c-Met  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: JC5148  
R:Aoki, S.; Takahashi, K.; Matsumoto, K.; Nakamura, T.  
J. Biochem. 120, 961-968, 1996  
A:Title: Molecular cloning of the Xenopus c-met/hepatocyte growth factor receptor and its  
A:Accession: JC5148  
A:Molecule type: mRNA  
A:Residues: 1-1375 <AOK>  
A:Cross-references: GB:AB027411; NID:g4877403; PIDN:BAA77764.1; PID:g4877404  
C:Comment: This protein is involved in early multiple organogenesis in Xenopus embryos.  
C:Genetics:  
A:Gene: c-met  
C:Superfamily: hepatocyte growth factor receptor; protein kinase homology  
C:Keywords: ATP; glycoprotein; growth factor receptor; phosphoprotein; phosphotransferase  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:1070-1338/Domain: protein kinase homology <KIN>  
F:1078-1086/Region: protein kinase ATP-binding motif  
Query Match 6.5%; Score 652; DB 1; Length 1375;  
Best Local Similarity 20.6%; Pred. No. 2e-33;



A:Accession: I57632  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 363-1009 <RES>  
A:Cross-references: GB:M5325; NID:g187531; PID:AAA59585.1; PID:g187532  
C:Comment: The receptor is a dimer of disulfide-bonded 50K alpha and 145K beta chains th  
C:Genetics:  
A:Gene: GDB:MET  
A:Cross-references: GDB:120178; OMIM:164860  
A:Map position: 7q31-7q31  
C:Superfamily: hepatocyte growth factor receptor; protein kinase homology  
C:Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferase;  
F:1-24/Domain: signal sequence #status predicted <IG>  
F:25-303/Product: hepatocyte growth factor receptor alpha chain #status predicted <ALP>  
F:308-1390/Product: hepatocyte growth factor receptor beta chain #status predicted <BET>  
F:933-955/Domain: transmembrane #status predicted <TM>  
F:1076-1344/Domain: protein kinase homology <KIN>  
F:1084-1092/Region: protein kinase ATP-binding motif  
F:45,106,149,202,399,405,635,785,930/Binding site: carbohydrate (Asn) (covalent) #status  
F:1110/Active site: Lys #status experimental  
F:1235/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status experim  
Query Match 6.5%; Score 651; DB 1; Length 1390;  
Best Local Similarity 19.9%; Pred. No. 2.4e-33;  
Matches 364; Conservative 266; Mismatches 557; Indels 640; Gaps 84;  
QY 62 HYLGVNRYIKLS-SDLKVLVTHETGDEDPKCYPRIVQTCNEPLTTI-----NNVN 115  
DB 61 HIFLGATNYIYLNEDLQKVAEYKTPGVLEHPDCCFP---CQDCSSKANLSGGVWKNIN 117  
QY 116 KMLIID-YKENRLIACSLYQGICIKLLRLLEDLFXLGFYHKKHYLSGVN-----ES 166  
DB 118 MALVVDIYDDQLISCSVNRGTQ-----RHVF---PHNHTADIQSEVHCIFSPQIEEP 169  
QY 167 GSVFGVIVS-----YNNLDK--LFIATVADGKPEYFP-----TISRRKLTKNSEADG 212  
DB 170 SQCPDCVVSALGAKVLSVXKDRFINFVGNNTINS--SYFPDPLHSISVRRLKETK--DG 225  
QY 213 MEAVVFHDEFVAMIKIPSDTFITIPDF---DIYVYVGSFGNVFVFLTLQPEMWSPPG 268  
DB 226 -FMFLTQDSVI-----DVLPERDSYPIKYVHAFESNNFIYFLTVORE-----267  
QY 269 STTKQVYTSKLVRLCKEDTAFNSYVEVPIGC-----ERSGVE--YRLLOAAYLSKAG 319  
DB 268 -TLDAQTFHTRIIRFCINGSLSHYMEMPLECILTEKRRKSTKKEVFENILQAAVYVSKPG 326  
QY 320 AVLGRTLGVHPDDLLFTVFSKQGRKKMSLDLSALCIFILKQINDRIKERLQSCYRGE 379  
DB 327 AQLARQIGASLNDLIFGVFAQSPDSABPMDRSAMCAFFIKYVNDFFNK-----376  
QY 380 TLDLAWLKVDIPICSSALLTIDDFGLDMNAPLGSDMVRGIPVFTEDRDRMTSVIAYV 439  
DB 377 -----IV 378  
QY 440 YKNSHLAFVGTSGKIKIRVDPGRGNALQVETVQVVDPGVPLRDMAFSKDHQIYINSE 499  
DB 379 NKNV-----RCLQH-----FYCPNHE-----395  
QY 500 RQLTRVPVSCGYQSCGECVLGSDPHCGWCVLHNTCTRKER-CERSKEPER--PASEMK 556  
DB 396 -----HC-----FNRTLNRSSGCEARDEYRTEFTALQ 425  
QY 557 QCVRLTVHPNNISVSQYNVLIV--LETYNVPELSAGVNCCTFEDISEMDGLVVGNIQCY 614  
DB 426 RV-----DLFWQPSFVLLTSIFIKGDL-----TIANLTSGRFQWVVSRS 471  
QY 615 PAAKEVPRIITENGHHVVOQLKSKETGTMFTASTSFVYNCNVHNSCLSCVSESPYRCHW 674  
DB 472 PSTPHVNFLL-----DSHPV-----SP-----488  
QY 675 CKYRHVCHTDPKTCSFQEGRVKLPEDCQQLLRVDKILVVPVVIKPIITLAKNLPQPSGQ 734

DB 489 -----EVIVEHTL-----NQ 498  
QY 735 RGYECILNIQSGEORVPALRFNSSVQCNTSYSEYEGMEINNLPLVELTVVWNG-----HFN 790  
DB 499 NGYTLVIT-----GKIKTIPL-----NGLGCRHF- 523  
QY 791 IDNPAQNKVHLKGMARESCGLCKADDDPACGWCQGGPGOCTLRQHPAESQWLELSG 850  
DB 524 -----QSCQCLSAFPFVQCWCCH--DKCVRSEEC-----LSG 554  
QY 851 AKSK--CTNPRITFIIPVTGREGTKTIRGENLGLFR-----DIASHVKVAGVECS 903  
DB 555 TWTOQICL-PAIYKVPFNSAPLEGTRLTICGWDG--FRNRKFDLKKTRVLLNESC 611  
QY 904 LVDGYIPAEQIVCEMGEAKSQHAGFVEICVAVCRPEFMASSQYLYPMTLTLSLKP 963  
DB 612 LTLSESTMTNLKCTVG--PAMNKH-FNMSIISNGHGTQYSTFSY-VDPVITSISPKY 666  
QY 964 GPMSSGGTQVITITGNLNAAGNVVMPGKQCLFHRHRSPIYVCNTTSSDEVLEMKV 1023  
DB 667 GPMAGGTLLTLTGNLNSGSRHISIGGKCTLKSVSNSILECYTPAQTIISTEFAVK 726  
QY 1024 DRAKIHODLVQYVEDPTIVRIEPMWSIVSGNTPIAVWGTHTLDLQNPQ--IRAKH 1081  
DB 727 DLAN-RETSIFSYPREDPLIVYEHPTKSPISGSTITGVKNLSVSVPRMWINVHEA 785  
QY 1082 HINICEVLNATE-MTCQAPALALGPDHQSILTERPEEFGFILDNVQSLIILNKTN 1140  
DB 786 FTVACQHRNSEIICCTPSSL--QQLNLQPLTKAF-FMLDGLSKYF-----DLIYV 837  
QY 1141 NPVEARPGSGILELKGTPPIILKGNLIPVAGNVKNTVLYGKPC-TVTVDVOL 1199  
DB 838 NPVEKPFKPMISGMNENVLEIRGNDIDPEAVKEV-----LKVGNKSCENIHLSE 892  
QY 1200 LCESPN-LIGHKVMARVGMGEYSP-----GMVYIAPDSPLS-LPAIVSIAGV 1249  
DB 893 LCTVPNDLLKLNSEL-----NIWQQAISSTVLGVQPDQNFCTGLIAGVSI 948  
QY 1250 IIFIVAVLIAYKRKRESDDLTKRLQMDQNLRESVALECKEFAELQTDIHELTD 1309  
DB 949 LGPFLWLKRRQIKDGLSELVRYDVAHTPHLDRLVSARSVSPTTMEVSN-----998  
QY 1310 AGIFFLDYRTYTMVLPGLPIEDHPVLRDLEVPYQERKEVGLKLFQALINNKFLL 1369  
DB 999 ---ESVDYRATFPBEDQFNSSONGSCROVQP-----1027  
QY 1370 RTLESQRSFMRDRGNVASLITVLQSKLEYATDVLKOLLADLIDKNLESKNHPK 1429  
DB 1028 -----LTD---MSPILTSGDS-----DISSPLQNTVHIDLSALN-PELV 1063  
QY 1430 TESVAEKMLTNWTFLLYKFLKECAGEPLF-----SLFCAIKQOMEKGPID 1475  
DB 1064 --QAVQHVIVGPFSSLIHV--FNEVIGRGHFGCVHGTLLDNDGKKIHCAVKS 1113  
QY 1476 AIT--GEARYSLSDKLIHQIDYKTLVLSVSPDNANSPEVPVKILNCDTITQVKE 1533  
DB 1114 RITDIGEVSQFTEGIMKDFSHPNVLSLLGLCLRSEGSPLV-----VLPYMKHGD 1165  
QY 1534 DAIFKNVPCSHRPAADM--DLEWROGS-----GARMLODEDTITKIEND 1577  
DB 1166 RNFINRE--TINPTVKDILGFLQVAKGMKYLASKKFVHRDLAARNCLMDEKFTKV 1222  
QY 1578 WKRLNTLAHQVDPGVSVALVSKQVATYANVNSTVSRTSASKYENMIRYTGSP 1637  
DB 1223 -----FGLARDMYDK-----YYSVHNKTGAKLPV-KWMAL-----ES 1258  
QY 1638 TPMITPDLES-GVKMWHLVKNHEHGDQEGDGRGSKWVSEI-----YLTRLATK 1690  
DB 1259 KFTTKSDVMSFGVJLWELM-----TRGAPPYDPDVTFTDITVYLLOGERLLQ 1307  
QY 1691 FVDDLFTIFSTAH-----RGSALPLAIKYMFD-FLDEQADKHGHPHVRHTWKS 1741  
DB 1308 CPDPLYEVLKCHWPKAEMRPSFSELVSRISAIFSTFICE-----HYVHVNA 1360

```
QY 1742 CL-----PLRFW 1748
Db 1361 CVAPYPSLLSSEDNADDEVDTTPASFW 1387

RESULT 13
S01254
hepatocyte growth factor receptor precursor - mouse
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) met
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S01254; JH0115; A45453
R:Chan, A.M.L.; King, H.W.S.; Deakin, E.A.; Tempest, P.R.; Hilken, J.; Kroezen, V.; Edw
Oncogene 2, 593-599, 1988
A:Title: Characterization of the mouse met proto-oncogene.
A:Reference number: S01254; MUID:88262253; PMID:2838789
A:Accession: S01254
A:Molecule type: mRNA
A:Residues: 1-1379 <CHA>
A:Cross-references: EMBL:X00671; NID:953058; PIDN:CAA68680.1; PID:953059
R:Wilks, A.F.; Kurbah, R.R.; Hovens, C.M.; Ralph, S.J.
Gene 85, 67-74, 1989
A:Title: The application of the polymerase chain reaction to cloning members of the prob
A:Reference number: JH0112; MUID:90152381; PMID:2482828
A:Accession: JH0115
A:Molecule type: mRNA
A:Residues: 'I', 1200-1254, 'R', 1256-1260, 'T', 1262-1268 <WIL>
A:Experimental source: hemopoietic cell
A:Note: The authors translated the codon ACG for residue 1261 as Lys
R:Weidner, K.M.; Sachs, M.; Birchmeier, W.
J. Cell Biol. 121, 145-154, 1993
A:Title: The Met receptor tyrosine kinase transduces motility, proliferation, and morpho
A:Reference number: A45453; MUID:93209981; PMID:8384622
A:Accession: A45453
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 924-935 <WEI>
C:Genetics:
A:Gene: met
C:Superfamily: hepatocyte growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; glycoprotein; phosphoprotein; phosphotransferase;
F.1-24/Domain: signal sequence #status predicted <SIG>
F.25-929/Domain: extracellular #status predicted <EXT>
F.25-302/Product: hepatocyte growth factor receptor alpha chain #status predicted <ACH>
F.308-1379/Product: hepatocyte growth factor receptor beta chain #status predicted <BCH>
F.930-954/Domain: transmembrane #status predicted <TM>
F.955-1379/Domain: intracellular #status predicted <INT>
F.1074-1342/Domain: protein kinase homology <KIN>
F.1082-1090/Region: protein kinase ATP-binding motif
F.1108/Active site: Lys #status predicted
F.1233/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predict

Query Match 6.5%; Score 646.5; DB 1; Length 1379;
Best Local Similarity 20.7%; Pred. No. 4.7e-33;
Matches 282; Conservative 188; Mismatches 411; Indels 483; Gaps 54;

QY 62 HYLGNVRIYKLS-SDLKVLVTHETGPDENPKCYPPRIVOTCNEPLTT-----NNVN 115
Db 61 HYLGNATNVIYVNDKDLQKVSFEKTPVLEHPDCLPFCR---DCSSKANSSGGVWKNDIN 117
QY 116 KMLLD-YKENRLIACGSLYQIGICK--LRLLEDLPKLGPPYHKKHYLSGVNESG---- 167
Db 118 MALLVDYDDQLISGVSNRGTQQRHVLPPDNSADIQSEVH---CMSPRESGQCPC 174
QY 168 --SVFGIVISYNNLDD--KLFIATAVDGK--PEY-FPTISSRLTKNSBADGMFAYVFDH 220
Db 175 VVSALGAKVLLSEKDRFINFFVGVNTINSSYPGYSLSHISVRLKETQ--DG-FKFLTDQ 231
QY 221 EFVAMIKIPSTFTIIPDF-----DIYVYVSGSGNFVFLTIQPEWSPPGSTTKEQY 276
Db 232 SYI-----DVLPEFLDSYPIKYTHAPENSHFIYELTVQKE-----TLDAQTF 273
```

Db 988 SPTTMSVN-----ESVDYRATFPEDQPNSSQNGACRQVYP 1025

RESULT 14

T30813

plasma1ogen related growth factor receptor 2 - Fugu rubripes

C:Species: Fugu rubripes

C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 15-Sep-2000

C:Accession: T30813

R:Cottage, A.J.; Clark, M.; Hawker, K.; Umrانيا, Y.; Wheller, D.; Bishop, M.; Elgar, G.

FEBS Lett. 443, 370-374, 1999

A:Title: Three receptor genes for plasminogen related growth factors in the genome of th

A:Reference number: Z20880; MUID:99148833; PMID:10025966

A:Accession: T30813

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1375 <COT>

A:Cross-references: EMBL:AJ010348; NID:e1355080; PID:e1355082; PIDN:CAA09100.1

C:Genetics:

A:Gene: PRGFR2

A:Introns: 400/3; 459/3; 499/3; 556/3; 611/2; 655/3; 701/2; 761/2; 794/3; 870/3; 919/3;

C:Superfamily: hepatocyte growth factor receptor; protein kinase homology

Query Match 5.6%; Score 559; DB 2; Length 1375;

Best Local Similarity 21.0%; Pred. No. 2.2e-27;

Matches 306; Conservative 189; Mismatches 442; Indels 522; Gaps 64;

8 WTCLSHLMV-----GWSSTLL---TRQAPLSQKRSFVTRGEPAGFNH 53

4 WTALLTICWILQTASGQHACASGSRTLVDFVKYSLPHFQTSK-----PVQ---N 52

54 LVV--DERTGHVILGAVNRIVKSLSDKLVLVTHETGDEONKCYPRIVQTC----- 104

53 IAVNWSQOKEIYIGCONAKAVSGTLEEMVEWETGP-VGSPDC-----ETCKLCDVEA 105

105 --NEPLTTNNVNMK-----LLIDYKENRLIACGSLYOGICKLLRLEDFKLGEPYHKEH 158

106 DPEDPVDTSEVLVDPAIGLLPY---LYVCGSTQHGICFYDIESPEHAPOCLYKKQR 161

159 YLGSVNES-----GSVFGVIVSYNLD---KLFIATVADGK-PHYFTTISRKLTK 206

162 ----NSPTKPCDCLASPLGPKVSi---IEDGSTYTFVFVAAVDDRVQRYPRRSISVWRP 214

207 NSEADGMFAVVFHDEFVASMIKIPSDFTIIPD---FDIYVYVFGSGNFWYFLTLQPE 262

215 LSTEDG-----FEMWTDGLTVLPSLSYKIDYIHSFATKYVYFLSLQRE 260

263 MVSPPGSTTKREQVYTSKLVRCKEDTAFNSYVEVPIGC-----ERSGVYRLLQA 312

261 --NFSNSNSPLQ---TRLGLPLSIREVMYREVVLVLECHFNPKRRRRGDFRGIYVNGLOA 315

313 AYLSKAGAVLGRVLGVHPDDLLFTVFS---KGOKRMKSLDSEALCIFILKQINDRIK 368

316 AHFRAGKDLAEELRVEDQEDILYGVFAVNVNELGTQR-----NSALCAFFPLSKVNHAID 370

369 ERLQSCYRGECTDLAWLKVKDIPCCSALLATIDNFCGLDMNAPLGVSDMVRGIPVFTED 428

371 EGVACCRS-----GTQLSRGLGHFQ-- 392

429 RDRMTSVIAYVYVKNHSLAFVGTGSKLKKIRVDGPRGNALQYETVVQVDPGVLDMAPS 488

393 ----- 392

489 KDHEQLYIMSERQLTRVPVSGQYQSCGECGLSGDPHCGMCLVANTCTRKERCERSKEP 548

393 -----PLESCP-----HESSEDKTC-RSK-- 411

549 RRFASEMKQCVRLTVHPNNISVSQYNVLLVLETVNVPELSAGVNTCFDLSEMDGLVGN 608

412 -----PTLVAQPHYR---LDLFN-----RMRDVLV-- 434

609 QIQCYSPAAKEVPRIITENGDDHHVVQLQLKSKETGWTASTSFVYFVNSVHNSCLSCVES 668

435 -----TTVMVTTTGNHTL----- 447

669 PYRCHWKYHVCTHDPKTCFQGRVVKLPEDCPQLLRVDKILVPVEVIKPTILKAKNLP 728

448 -----GHFTSD-----GRI-----LQVILSLYRIVFANYSLG 476

729 PQSGQGRGYECILNIQSGEQRVPALRFNSSVQCONTSYSVEGMEINNLPVELTVVMNGH 788

477 -----DGEVSRTRAV-----YSEDWL-----LFVGVGNKM 500

789 FNTDNPANQKHLVKCAMRESCLCLKADDPFACGCGQCTLRQHCFAQBSOWLEL 848

501 FKVPSLPGGCAHF-----RTCSMCLMAPRFMNCGWS--GVCSRQHC--DMQW--- 545

849 SGAKSKCTNPRITEIIPVTGPREGKTIVIRGENGLGLEFRDIASHVKVAGVECSPLVDGY 908

546 --EKDSCA-PVITEFSPKVPVPGDTEVTLGWFQSPLRPAIINGKTHSIMVGSTVCIV 602

909 IP-----AEQIVCEMGEA-KPSQHAGFVEICVAVCRPFMARSS-----QIYYFMTLTL 956

603 LPEKSSSEKLVCRIQKNKPGQD---LKIVLKVHEGEVEGRYSIEGTAQIQGFSVEFSI 659

957 SDLKPSRGPMSGGTQVTTITGNLNASGNVVMFGKQPCLFHRS-----PSVIVCNTTSS 1011

660 TDIQPDYGPAPGGTFTVTLGRHLDGSGFQDVFVEKQCRILSVSGGTDASSIVCLPAVA 719

1012 DEVLEMKVSQVDRAKIHQDLVFOYVEDPTTVRIEPEWSIVSGNTPIAVWGTHLDLIQNP 1071

720 EDVGSVPVKLIDSFVTATQFFYKRNPNVITLVEPQCSLHSGSR-LVIQGNLDSAHKT 778

1072 QIRAKHGKKEHIN-----ICE-VLNATEMTCQAPALAGDPHOSDLTERPEEPGF 1120

779 VW-----EYVSKDAMLRLQOVQCNSTNATMECWAPAL---PEEMPD-----EKETQ 824

1121 ILDNVQSILLINKNTFTYPN-PVFEAFPGSGILELKEG-TPILILKGNLIPPVAGNVK 1178

825 IYIHMEBQLKVMSSRFYHGNVKVPIPPENEDRVLLKKEGENEVSLLHRLN--NTVSLCNK 882

1179 LNYTVLVGEKCTVTVDVQLLCESPNLI---GRHKVMARVGGMEYSPGMVYIAPDSPL 1234

883 INMT--VGVVDCAEVLNLTCTRIKGLVPTGGLPVKVSIGGEVYDVGTGIDDSNAA 940

1235 SLPAIVIAVAGLLIIFIVAVLIAYK-RKRESDLTLKRLQOMQDNLESRLVALECKEAF 1293

941 VSIATIVILSIVGALVIGAVMALAVMFHVRKXKCANIE-----HRLSTRUSLARSSAA 992

1294 AELQT--DIH--ELTSDLDGAG-----IPFLDYRTYM----- 1322

993 ADFSTGDRYRDVLSSTPSGSGMAFOGLLYAASSDHLISIPLLPRDSISMVSLNPDLEE 1052

1323 --RVLFPQ-----IEDHPVL 1335

1053 VKNVLIPADMLRIEDRQVI 1071

RESULT 15

T30811

hepatocyte growth factor receptor - Fugu rubripes

C:Species: Fugu rubripes

C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 15-Sep-2000

C:Accession: T30811

R:Cottage, A.J.; Clark, M.; Hawker, K.; Umrانيا, Y.; Wheller, D.; Bishop, M.; Elgar, G.

FEBS Lett. 443, 370-374, 1999

A:Title: Three receptor genes for plasminogen related growth factors in the genome of the

A:Reference number: Z20880; MUID:99148833; PMID:10025966

A:Accession: T30811

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1425 <COT>

A:Cross-references: EMBL:AJ009961; NID:e1354962; PID:e1354963; PIDN:CAA08917.1

C:Genetics:

A:Gene: PRGFR1

A: Introns: 407/3; 475/3; 518/3; 576/3; 629/2; 667/3; 713/2; 768/2; 801/3; 871/3; 921/3;  
 C: Superfamily: hepatocyte growth factor receptor; protein kinase homology

Query Match	5.6%	Score 556;	DB 2;	Length 1425;
Best Local Similarity	20.6%	Pred. No. 3.6e-27;		
Matches 372;	Conservative 223;	Mismatches 565;	Indels 648;	Gaps 83;
15	LLVMGSGSTLLTROPAPLQKQSFV-----TFRGE-PAEGFNHLVDERTGHIYLGAV 68			
11	LLLWALSAGHCQKSPENKLNLSVSEYELFTFAEPFIQ--NVVTLN-----GIIIVGAT 65			
69	NRIYKLSDLKVLVTHETGCPDBDNPKYPPRIVQTCNEPLTTT-----NNVKNMLLID 121			
66	NRIYALAPSLTKLSYRTGTPLLAN-----QTCGQKVANASSGGKRDNLNVALVVE 116			
122	-YKENRLTACGSLYOGICLKLLEDFLKGEPVHKKEHLYSGVSGSVFVTVSYNLT- 179			
117	NIYDGLFSCGADNGVCRHRLVEDVSLDEGRQK-----SVDELVCYCFIDUK 165			
180	DDKLFATATADGKPYFPTTISRKLTK--NSBADGMFAVYVHDEFVAS----- 225			
166	QDK-----GQPRSDVVVSPSGSVQVLNVSSEIMFFVGNSEIPGSGNVTGPTARPH 217			
226	-----MIKIPSDTFT-----IIP-----DFDIYVYVFGSGNFVYFLTLQPEMVSPPG 268			
218	MSLRKMTSQNGFTFFSNRSMYDLIPPLRGSYLYRYVYFSHSGPFTYFLTVQ----- 269			
269	SITKE-QVYTSKVLRLCKEDTAFNSVVEPIGC-----ERSGVE-----YRLLOAYLSK 317			
270	QVSKDSQYTHTRIVRMCSDDHDIRRYVEMPLEBICSTDKRRRRSMEDVKFNILQAAVTK 329			
318	AGA--VLGRTLGVHPDLDLFTVFGKQKRMKSLDESALCIFILKQINDRIKERLQSCV 375			
330	VGSDVELQRLBEGDDVLFARFARGKNSTEATPNSALCVMSLKLINSMFQWYQKC- 388			
376	REGGLDLAWLKVKDIPCSALLTTIDNFCGLDMNAPLGVSDMVRGIPVFTEDRDMTSV 435			
389	-----NTVDLY-----HFTGSDK----- 401			
436	IAYVYKNHSLAFVGTGSKLKKIRVDGPRGNALQYETVQVDPGVLRDMAFSKDEHLY 495			
402	----- 401			
496	IMSERQLTRVPVESCQYQSCGECILGSDPHCGWCVLHNTCTRKERCERSKPRPFASEM 555			
402	-----KSCYNVSSDDC-----DPHEG---IHEGKEGYRLQVTFQVORLEYWQ 442			
556	KQCVRLTVHPNNISVQYNVLVLETYNVPELSAGVNCFTFEDLSEMDGLVGVNQIOCYSP 615			
443	KVL-----TNTLVTSITWTV-----HGRAVG----- 464			
616	AAKEVPRIITENGDDHVVQQLKSKETGKTFASTSFVYNCVHNSCLCSVPSPYRCHWC 675			
465	-----YLTADGRH--IQV-----VF-----SRFASP----- 484			
676	KYRHVCTHDPKTCFQEGRVKLPEDCPQLLRVNBKILVPVEVIKPIITLAKNLPPQSGQR 735			
485	-----HV-----NIRLDSRPVSGSVLP----- 502			
736	GVECILNIQGSORVPALRPNSSVQCNSTSYSGMEINNLVELTVVNGHFNIDNPA 795			
503	-----GQDPSECALLTT-----GNKIKTVB-----LIGPG 528			
796	QNKVHLKCGAMRESGLCLKADPDFACGWCQPGGCTLRQHCPAQESQWLELSGAKSKC 855			
529	CD--HL-----TTCTSLVSRVTECGWCE--GRCTRANQCP--PSVM-----TQEYC 570			
856	TNPRITEIIPVTPREGGKTVTIRGENLGL-----FRDIASHVKVAGVECSPLVDGVIPA 911			
571	T-PVTVKVPFTSGPIRGSTVTCMGRNFGDKTESFKASLVTVEVAGVPCKLRSQYASR 629			
912	-EQIVCE---MGEAKPSQHAGFVEICVAVCRPBFMARSSQLYFFMTLTLSDLKPSRGPMS 967			

Search completed: June 24, 2003, 19:01:00  
 Job time : 80 secs

630	WTBIQCSMPFSGNFTPSGHT-----VKVTSGHKIA-TIEGFTFVDPVSEIFPTFGPKS 682	Db
968	GGTOVTITGNLGNAGSNVVMFGKQCLFHRHRSYIVCTNTSSDEVLEMKVSQVQVDRAK 1027	Qy
683	GNTMLTIRGAFLDTGNKREVTVGKAACKIOSLSATMLTCKTPHAPVSKQPVRLTVDSVA 742	Db
1028	IHQDLVFQVVEDPTIVRIPEWSIVSGNTPIAVGTHDLIDIONPQIRAKHGKKE---HIN 1084	Qy
743	RDAPVLYTHQDPIIISIQPSRFSVSGCTVAAAGHLFQSGLOPQMWLTTCQDAEVEFHS 802	Db
1085	ICEVLNATENTCOAPALALGPDHQSDLTERP---EEFGFILDNVQSLILLKNTNFTYYPN 1141	Qy
803	CVYGENRTSIQCTTPSLA-----KLAQPPVVTKVAFLDGYMT---EQWOLLYVED 851	Db
1142	PVPEAFPGSGILEKPGTPII-LKGNKL-----IPVAGGNVKLNTVILVGEK-PCTV- 1192	Qy
852	PLFQ--DPK--LTSKDNKSIVELKGDMDREAMKCQVLTVSNHSCESLTLVGNTLECTVP 907	Db
1193	-----TVSDVQLLCSBNLIGRHKNVAVRGMEYSPGVYIAPDPLSLPAIVSIIVAG 1246	Qy
908	TELQTTTSKELQVEWRQADSI-RHL-----GKVTLAQEQDYTGILVGMCVS- 953	Db
1247	GLLIIFIVAVLIAKRSRESLDT-----LKRLOMQMDNLESRLVALSCKEAPA 1294	Qy
954	-----LTLGLLGLLWKKKHHIDLSVYDGRGHIQHLDRL-----ANA 994	Db
1295	ELQTDIHELTSDLGAGIPPLDYRTYTMRVLPFGIEDHPVLRDLDEVPYRQERVEKGLKL 1354	Qy
995	RSVSPTEWVSH-----ESVDYRTNL-----LEDO-----GSDRETCEAGPPI 1033	Db
1355	P---AQLNNKVFLLSFIRTLQSRSFMSDRGNVASLIMTVLQSKLEYATDVLKQLLAD 1411	Qy
1034	YCGNGELLSPLVALG-----SGLGLGMEGELVSPLLMAPVHIDPSCL----- 1076	Db
1412	LIDKNLESKNHPKLL--LRRTESVAEKMLTN-----WFTFLLYKFLKECAGEPLESL 1461	Qy
1077	-----HPDLLTEVQHVVIAREQLLHLNQVIGRHFHGCVFHGTLLLEPDGQ---KQ 1123	Db
1462	FCAIKQOMEKGPIDAITG--EARYSLSSEDKLIRQIDYK-----TLVLSCVSPDNANSPE 1514	Qy
1124	HCAVKS-----LNRITDLEEVQFLKEGLIMK---DFSHPNVLSLIGLICLPPEGS----- 1170	Db
1515	VPVKILNCDTITQVKEKILDAIFKNVPCSHRPKAADM--DLEWRQGS----- 1559	Qy
1171	-PLMWLPYMKHGLRNPFRDE-----SHNPTVKDLMGFLQVARGMEYLASKKFVHRD 1222	Db
1560	-GARMLODEDIITKTN-----DVKRLNTLAHVQVDPGS 1593	Qy
1223	FAARNCLMDESYTVKADFGIARDVVDKEYYSVHNKSGVKLPVKWMALESQTHKFTSKS 1282	Db
1594	VV-----ALYSKQVTAYNVNNSTVSRSTSASKYENMIRYTGSPDSLSRSTPMITPDL 1645	Qy
1283	DVWSFGVLLWELMELTRAPPYSDVNSPDIT-VFLQGRRLQPEFCPDAL--YTVMI----- 1335	Db
1646	ESGKMMH 1653	Qy
1336	-----ECWH 1339	Db





FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 548 548 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 738 738 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 746 746 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 1009 1009 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 1036 1036 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 1073 1073 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 1115 1115 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
FT CARBOHYD 1162 1162 N-LINKED (GLCNAC. .) (POTENTIAL. .) (POTENTIAL. .)  
SQ SEQUENCE 1871 AA; 207661 MW; 28420CEBDB22E9CE CRC64;  
Query Match 59.8%; Score 5988; DB 1; Length 1871;  
Best Local Similarity 60.0%; Pred. No. 0;  
Matches 1133; Conservative 293; Mismatches 437; Indels 26; Gaps 13;  
QY 10 CLLSHLMVWGSGSTLTROPAPLSQKORSQVFTFRGEPAGFNHLVVDERTGHIYLGAVN 69  
DB 5 CLLL-LFLAVGGA--LGNRPF-----RAFVV-----TDITLTHLAVHRTGVEFVGAVN 51  
QY 70 RIYKLSDLVIVTHETGPDNDPKYPPRIVOTCNEPLTTNNVKNMLLDYKENELIA 129  
DB 52 RVFKLAPNLTELRAHVTGPFVEDNARCPYPPSMRVCAHRLAPVDNINKLLLDIYAARLVA 111  
QY 130 CGSLYOGICLLLEDLFKLGEPPHKEHYLSGVNCSGVSFGVIVSYNSNLDLKFIAFATV 189  
DB 112 CGSIWQICQFLRLDDLFKLGEPPHKEHYLSGAEQEPDSMAGVIVEOGQSKLFVGTAV 171  
QY 190 DGRPEYPTTSRRLTKNSADGMFAVYFHFVASMIKIPSDFTTIPDPIIYVYVGS 249  
DB 172 DGRSEYPTTSRRLTKNSADGMFAVYFHFVASMIKIPSDFTTIPDPIIYVYVGS 231  
QY 250 SGNFVYLTLOPEMVSPGSTTKQOVVTSKLVCKEDTAFNSVVEPIGERSGVEYRL 309  
DB 232 SASFVYLTLOPEMVSPGSTTKQOVVTSKLVCKEDTAFNSVVEPIGERSGVEYRL 291  
QY 310 LOAYLSKAGAVLGRITLGVHPDLDLLFTVPSKGQRKMSLDESALCIFILQNDRIKE 369  
DB 292 VQSHLAKPGLLAQALGVPADEVDLFTFSQGNKRPQITLCLFSLNINAHRR 351  
QY 370 RLQCYRGEGLDLAWLKVDIIPCSALLTDDNFCGLDMNAPIGVDMVRGIPVFTEDR 429  
DB 352 RIQCYRGEGLTALPWLNLKLPCLINTPQXNGNFCGLVNLQPLGLHVEIGLEPLLDST 411  
QY 430 DRMTSVIYVYKXSLAFVGTGSKLIRVVDGRGNALQVETVQVDPGVPLDMAPSK 489  
DB 412 DGNASVYIYVYKXSLAFVGTGSKLIRVVDGRGNALQVETVQVDPGVPLDMAPSK 470  
QY 490 DHEQYIMSERQLTRVPVESCQYQSCQCECLGSDPHCGWCVLHNTCTRKERCERSKEPR 549  
DB 471 DHRHILLSEKQVSQLPVETCEQYQSCAACLGSDPHCGWCVLHNRCCREGACILGASAPH 530  
QY 550 RPASEMKQCVRLTVHPNNISVQVNLVLETVNVPFELSAGVNCFTFEDLSEMDG-LVVG 608  
DB 531 GFABELSKCVQVRVRPNVSVTSQVQLTVTLHNPVLSAGVSCAFEAANEAVLLPSG 590  
QY 609 QIQCYSPAACEVPRILITENGHHVVLQKSKETGTEASTSFVYNCVSHNCSCLSVES 668  
DB 591 ELLCPFSLQELRALTRGHGATVRLQLLSEKTVRGVRFAGADFPYNCVSVLQSCMSCVGS 650  
QY 669 PYRCHWKYRHTVTHDPKTCFSQEGRVKLPBDCPQLLRVDKILVPEVVIKPIITLAKNLP 728  
DB 651 PYPCHWKYRHTVTHDPKTCFSQEGRVKLPBDCPQLLRVDKILVPEVVIKPIITLAKNLP 710  
QY 729 QPQSGQRYECLINIQSGEQRVPALRFNSSVQOQNTSYSEHGEINNLPELTVWNGH 788  
DB 711 QPQSGQRYECLINIQSGEQRVPALRFNSSVQOQNTSYSEHGEINNLPELTVWNGH 770  
QY 789 FNIDNPAQNVHLKCGAMRESGLCLKADPDPACGQCGQPCQCTLRQHCPAQESOWLEL 848  
DB 771 FPDKPSFRALLKCAQORSCGLCLKADPDPACGQCGQPCQCTLRQHCPAQESOWLEL 830  
QY 849 SGAKSCKTNPRITEIIPVTGREGGKTIVIRGENLGLFRDIASHVKVAGVECSPLVDGY 908

DB 831 SOXGTRCSHPRIITQIHPLVGPKEGGTRVTVIGENGLLSREVG--LRVAGVRCNSIPAEY 888  
QY 909 IPAEQIVCEGEEA-KPSQAGAFVEICVAVCEPFWASSQLYFMTLLSLDKPSRGPM 967  
DB 989 ISABRIIVCEMEESLVPSPFPGPVELCDGCSADPRTQSEQVIFVTPPTDQVSPRGPM 948  
QY 968 GGTQVITITGNLNAAGSNVVMFGKQPCFLFRRSPSYIVC-NTTSSDEVLKMSVQVDR 1026  
DB 949 GGTFLTITSSSLDAGSRVTVTRDSECOFVERDAKAIVCISPLSTLGPQAFITLAIDRA 1008  
QY 1027 KIHQ-DLVQFQVEDPTTIRIEPEWSIVSGNTPIAVWGTHLDLIQNPQIRAKHGGEHINI 1085  
DB 1009 NISSPGLIYVTTQDPTVTRLEPTWSIIINGSFAITVSGTHLLTVQEPVRARKGIEETNT 1068  
QY 1086 CEVLNATEMTCQAPALALGPDHQSDLTERPEFGFILDNVOSLLILNKTNTFYYPNPVE 1145  
DB 1069 CQVINDTAMLCAPGIFLGRPQRAQEGHPDEFGFLDHQVQTARSLNRSSFFIYDPSE 1128  
QY 1146 AFGPSGILELKPCTPIILKGRNLIPPVAGGNVKNLYTLVGEKPCCTVTVSDVOLLCESPN 1205  
DB 1129 PLGSGVLVDKPGSHVVLKGNLIPAAAGSS-RLNVTVLIGQPCSLTVSDTQLLCDSPS 1187  
QY 1206 LIGRHKVMARVGMVSPGMVYIAPDSPLSIPAIVSIAVAGLLIIFIVAVLIAKXKR 1265  
DB 1188 QTGRQPVVVLVGGLEFWLGTLHISAERALTLPAMMGLAAGGLLALITAVLVAYKRTQ 1247  
QY 1266 ESDTLTKRLQOMONLESVALECKEAPAELOTDIHELTSLDGAGIPFLDYRTYTMVRL 1325  
DB 1248 DADTLTKRLQOMONLESVALECKEAPAELOTDIHELTSLDGAGIPFLDYRTYTMVRL 1307  
QY 1326 PFGIEDHVPVLRDLVPGVQRBVEKGLKLAQLINNKVFLSFIITLSQSFMSMRDGN 1385  
DB 1308 PFGIEAHPVLKELDTP---PNVEKALRLFGQLLSHRAFLVTFIHTLEAQSSFSMRDGT 1363  
QY 1386 VASLIMTVLOSKLEVATDVLQALLADLIDKNLESKNHPKLLILRRTESVAEKMLTNWFTL 1445  
DB 1364 VASLIMTVLOSKLEVATDVLQALLADLIDKNLESKNHPKLLILRRTESVAEKMLTNWFTL 1423  
QY 1446 LYKFLKECAGPLSLFCAIKQOMEKGPIDAITGEARYSLSEDKLIRQOQIDYKTLVLSCV 1505  
DB 1424 LHKFLKECAGPLSLFCAIKQOMEKGPIDAITGEARYSLSEDKLIRQOQIDYKTLVLSCV 1483  
QY 1506 SPDNANSEPEVPVKILNCDDTITQVKEKILDAIFKNVPCSHRPAKADMDLEWQSGGARMIL 1565  
DB 1484 CPENEGSAQVFPVKVNLCDSTIQAOKDLDTYKIGIPYSORPKAEDMDLEWQSGGARMIL 1543  
QY 1566 QDEDTITTKIENDWKELNLAHVQVDPGVSVALVSKQVTAYNVAVNSSTVSRTSASKYENMI 1625  
DB 1544 QDEDTITTKIENDWKELNLAHVQVDPGVSVALVSKQVTAYNVAVNSSTVSRTSASKYENMI 1602  
QY 1626 RYTGSPDSLRSRTPMITPDLESQVKNVHLVKNHEHGDQKEGDRGSKVMSVSEIYLRLLATK 1685  
DB 1603 RTASPSDSLRSRTPMITPDLESQVKNVHLVKNHEHGDQKEGDRGSKVMSVSEIYLRLLATK 1662  
QY 1686 GTLQKFVDDLFTETIPSTAHRSALPLAIKYMPTDFDEQADKHGIDHPHVRHTWKSCLPL 1745  
DB 1663 GTLQKFVDDLFTETIPSTAHRSALPLAIKYMPTDFDEQADKHGIDHPHVRHTWKSCLPL 1722  
QY 1746 RFWNMINKPQVFDIHKNSITDACLIVVAQTFMDSCSTSEHRLGKDSFNSKLLIYAKDIP 1805  
DB 1723 RFWNMINKPQVFDIHKNSITDACLIVVAQTFMDSCSTSEHRLGKDSFNSKLLIYAKDIP 1782  
QY 1806 SYKNWERYYSIDIGKWPASISQDMNAYLAEOQSRMHNENFTMSALSEIFSYVGYKSEIL 1865  
DB 1783 NYKSWERYYSIDIGKWPASISQDMNAYLAEOQSRMHNENFTMSALSEIFSYVGYKSEIL 1842  
QY 1866 GFLDHDQCGKOKLAYKLEQVITILMSLDS 1894  
DB 1843 TALDRDASCRKHLRQKLEQIISLVSSDS 1871

MET\_HUMAN  
ID MET\_HUMAN STANDARD; PRT; 1390 AA.  
AC P08581; Q9UPL8; O60366;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hepatocyte growth factor receptor precursor (BC 2.7.1.112) (Met proto-oncogene tyrosine kinase) (c-met) (HGF receptor) (HGF-SF receptor).  
GN MET.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Giordano S.;  
RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX MEDLINE=87317655; PubMed=2819873;  
RA Park M., Dean M., Kaul K., Braun M.J., Gonda M.A., Vande Woude G.;  
RT "Sequence of MET protooncogene cDNA has features characteristic of the tyrosine kinase family of growth-factor receptors.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:6379-6383(1987).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RA Pauley A., Andrews S., Lacy M., Holmes A., Harper M.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 1010-1390 FROM N.A.  
RX MEDLINE=88143699; PubMed=3325883;  
RA Chan A.M.L., King H.W.S., Tempest P.R., Deakin E.A., Cooper C.S., Brookes P.;  
RT "Primary structure of the met protein tyrosine kinase domain.";  
RL Oncogene 1:229-233(1987).  
RN [5]  
RP SEQUENCE OF 1206-1264 FROM N.A.  
RX MEDLINE=94067791; PubMed=8247543;  
RA Lee S.-T., Strunk K.M., Spritz R.A.;  
RT "A survey of protein tyrosine kinase mRNAs expressed in normal human melanocytes.";  
RL Oncogene 8:3403-3410(1993).  
RN [6]  
RP SEQUENCE OF 1267-1390 FROM N.A.  
RX MEDLINE=86065462; PubMed=4069211;  
RA Dean M., Park M., le Beau M.M., Robins T.S., Diaz M.O., Rowley J.D., Blair D.G., Vande Woude G.F.;  
RT "The human met oncogene is related to the tyrosine kinase oncogenes.";  
RL Nature 318:385-388(1985).  
RN [7]  
RP FUNCTION.  
RX MEDLINE=91118019; PubMed=1846706;  
RA Bottaro D.P., Rubin J.S., Faletto D.L., Chan A.M.-L., Knicek T.E., Vande Woude G.F., Aaronson S.A.;  
RT "Identification of the hepatocyte growth factor receptor as the c-met proto-oncogene product.";  
RL Science 251:802-804(1991).  
RN [8]  
RP PHOSPHORYLATION OF TYR-1235.  
RX MEDLINE=92011756; PubMed=1655790;  
RA Ferracini R., Longati P., Naldini L., Vigna E., Comoglio P.M.;  
RT "Identification of the major autophosphorylation site of the Met/hepatocyte growth factor receptor tyrosine kinase.";  
RL J. Biol. Chem. 266:19558-19564(1991).  
RN [9]  
RP VARIANTS HPRC, AND VARIANT VAL-320.  
RX MEDLINE=97285124; PubMed=9140397;  
RA Schmidt L., Dutt F.-M., Chen F., Kishida T., Glenn G., Choyke P., Scher S.W., Zhang Z., Lubensky I., Dean M., Allikmets R., Chidambaram A., Bergerheim U.R., Feltis J.T., Casadevall C., Zamarron A., Bernues M., Richard S., Lips C.J.M., Walther M.M., Tsui L.-C., Geil L., Orcutt M.L., Stackhouse T., Lipan J., Slife L., Brauch H., Decker J., Niehans G., Hughson M.D., Moch H., Storkel S.,

RA Lerman M.I., Linehan W.M., Zbar B.;  
RT "Germline and somatic mutations in the tyrosine kinase domain of the Met proto-oncogene in papillary renal carcinomas.";  
RL Nat. Genet. 16:68-73(1997).  
CC -1- FUNCTION: RECEPTOR FOR HEPATOCYTE GROWTH FACTOR. HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC -1- SUBUNIT: HETERODIMER FORMED OF AN ALPHA CHAIN (50 kDa) AND A BETA CHAIN (145 kDa) WHICH ARE DISULFIDE LINKED.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- DISEASE: ACTIVATION OF MET AFTER REARRANGEMENT WITH THE TPR GENE PRODUCES AN ONCOGENIC PROTEIN.  
CC -1- DISEASE: DEFECTS IN MET ARE THE CAUSE OF HEREDITARY PAPILLARY RENAL CARCINOMA (HPRC). HPRC IS A FORM OF INHERITED KIDNEY CANCER CHARACTERIZED BY A PREDISPOSITION TO DEVELOP MULTIPLE, BILATERAL PAPILLARY RENAL TUMORS. THE PATTERN OF INHERITANCE IS CONSISTENT WITH AUTOSOMAL DOMINANT TRANSMISSION WITH REDUCED PENETRANCE.  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M35074; AAA59590.1; --  
CC EMBL; X54559; CAB56793.1; --  
CC EMBL; AC002080; AAB54047.1; --  
CC EMBL; AC004416; AAC06178.1; --  
CC EMBL; AC004416; AAC06179.1; --  
CC EMBL; J02958; AAA59591.1; --  
CC PIR; A40175; TVHUME.  
CC HSP; P11362; IFGK.  
CC Genew; HGNC:7029; MET.  
CC MIM; 164860; --  
CC InterPro; IPR000719; Euk\_pkinase.  
CC InterPro; IPR002909; IPT\_TIG.  
CC InterPro; IPR003659; plexin-like.  
CC InterPro; IPR002165; plexin\_repeat.  
CC InterPro; IPR001627; Sema.  
CC InterPro; IPR001245; Tyr\_pkinase.  
CC Pfam; PF00069; pkinase; 1.  
CC Pfam; PF01403; Sema; 1.  
CC Pfam; PF01437; PSI; 1.  
CC Pfam; PF01833; TIG; 3.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom; PD000001; Euk\_pkinase; 1.  
CC SMART; SM00429; IPT; 4.  
CC SMART; SM00423; PSI; 1.  
CC SMART; SM00219; TyrK; 1.  
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
CC Transferase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding; Receptor; Transmembrane; Glycoprotein; Phosphorylation; Signal;  
KW Alternative splicing; Chromosomal translocation; Disease mutation; Polymorphism.  
FT SIGNAL 1 24  
FT CHAIN 25 1390 HEPATOCYTE GROWTH FACTOR RECEPTOR.  
FT DOMAIN 25 932 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 933 955 POTENTIAL.  
FT DOMAIN 956 1390 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 55 500 SEMA.  
FT DOMAIN 1078 1345 PROTEIN KINASE.  
FT NP\_BIND 1084 1092 ATP (BY SIMILARITY).  
FT BINDING 1110 1110 ATP (BY SIMILARITY).  
FT ACT\_SITE 1204 1204 BY SIMILARITY.



Db 1114 RTDGEVSQFTEGIIMKDFGHPNVLISLGLCLARSEGSPV-----VLPMYKHGDL 1165  
QY 1534 DAIFKNVPCSHRPKAADM--DLEWROGS-----GARMLODEIDITTKIEND 1577  
Db 1166 RNFINRE--THNPTVKDLIGFLOVAKGMKYLASKKFFVRDLAARCMLEDEFTVKVAD- 1222  
QY 1578 WKRLNTLAHYQVPGDSVALVSKQVYAYNAVNNSTVSRTSASKYENMIRYTGSPDSLSR 1637  
Db 1223 -----FGLARDMYDKE--YYSVHNKTGAKLPV-KWAL-----ESLQTO 1258  
QY 1638 TPMITPDLES-CVKMHLVKNHEHGDQEGDRGSKWVSEI-----YLTRLIATKGTLOK-- 1690  
Db 1259 KFTTKSDVMSFGVLMELM-----TGAPPYDVNTFDITVYLLQRRLLQPEY 1307  
QY 1691 FVDLLFEITFAH-----RGSALPLAIKMYFD-FLDEQADKHGHPHVRHTWKS-N 1741  
Db 1308 CPDPLVEVWLKWHKPAEMRPFSELSVRSIAFSTFICE-----HYVHNATVNVK 1360  
QY 1742 CL-----PLRFW 1748  
Db 1361 CVAPYPSLLSSEDNADDEVTRPASFW 1387

RESULT 3  
MET\_MOUSE  
ID MET\_MOUSE STANDARD; PRT; 1379 AA.  
AC P16056; Q62125;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hepatocyte growth factor receptor precursor (EC 2.7.1.112) (Met proto-  
oncogene tyrosine kinase) (c-met) (HGF receptor) (HGF-SF receptor).  
GN MET.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88262253; PubMed=2838789;  
RA Chan A.M.L., King H.W.S., Deakin E.A., Tempest P.R., Hilkens J.,  
Kroezien V., Edwards D.R., Wills A.J., Brookes P., Cooper C.S.;  
RT "Characterization of the mouse met proto-oncogene.";  
RL Oncogene 2:593-599 (1988).  
RN [2]  
RP SEQUENCE OF 1199-1270 FROM N.A.  
RX MEDLINE=90152381; PubMed=2482828;  
RA Wilks A.F., Kurban R.R., Hovens C.M., Ralph S.J.;  
RT "The application of the polymerase chain reaction to cloning members  
of the protein tyrosine kinase family.";  
RL Gene 85:67-74 (1989).  
CC -!- FUNCTION: RECEPTOR FOR HEPATOCYTE GROWTH FACTOR. HAS A TYROSINE-  
CC -!- PROTEIN KINASE ACTIVITY.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBUNIT: HETERODIMER FORMED OF AN ALPHA CHAIN (50 kDa) AND A BETA  
CC CHAIN (145 kDa) WHICH ARE DISULFIDE LINKED.  
CC -!- SURCELLULAR LOCATION: TYPE I membrane protein.  
CC -!- DISEASE: ACTIVATION OF MET AFTER REARRANGEMENT WITH THE TPR  
CC (TRANSLOCATED PROMOTER) LOCUS OF CHROMOSOME 1 PRODUCES AN  
CC ONCOGENIC PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
CC -!- SIMILARITY: CONTAINS 1 SEMA DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Y00671; CAA68680.1; -.

DR EMBL; M33424; AAA40015.1; -.  
DR PIR; S01254; S01254.  
DR HSSP; P11362; 1FGK.  
DR MGD; MGI:36569; Met.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002909; IPT\_TIG.  
DR InterPro; IPR003659; plexin-like.  
DR InterPro; IPR002165; plexin\_repeat.  
DR InterPro; IPR001627; Sema.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF01403; Sema; 1.  
DR Pfam; PF01437; PSI; 1.  
DR Pfam; PF01833; TIG; 3.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00429; IPT; 4.  
DR SMART; SM00423; PSI; 1.  
DR SMART; SM00219; TyrKC; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;  
KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 1379 HEPATOCYTE GROWTH FACTOR RECEPTOR.  
FT DOMAIN 25 931 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 932 954 POTENTIAL.  
FT DOMAIN 955 1379 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 55 499 SEMA.  
FT DOMAIN 1076 1343 PROTEIN KINASE.  
FT SITE 306 307 CLEAVAGE (POTENTIAL).  
FT NP\_BIND 1082 1090 ATP (BY SIMILARITY).  
FT BINDING 1108 1108 ATP (BY SIMILARITY).  
FT ACT\_SITE 1202 1202 BY SIMILARITY.  
FT MOD\_RES 1233 1233 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 106 106 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 201 201 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 357 357 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 398 398 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 404 404 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 606 606 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 634 634 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 784 784 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 878 878 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CONFLICT 1199 1199 V -> I (IN REF. 2).  
FT CONFLICT 1255 1255 T -> R (IN REF. 2).  
FT CONFLICT 1261 1261 K -> T (IN REF. 2).  
FT CONFLICT 1269 1270 VL -> IP (IN REF. 2).  
SQ SEQUENCE 1379 AA; 153548 MW; FC5CC87FDD8ADEB8 CRC64;  
Query Match 6.5%; Score 646.5; DB 1; Length 1379;  
Best Local Similarity 20.7%; Pred. No. 1.8e-33;  
Matches 282; Conservative 188; Mismatches 411; Indels 483; Gaps 54;  
QY 62 HYILGAVNRIYKLS-SDLKVLVTHETGPDNDPKYPPRIVQTCNEPLTT-----NNVN 115  
Db 61 HYILGATNIIYVUNDKDLQKVSEFKTGPVLEHDLCLPCR---DCSSKANSSGGVWKONIN 117  
QY 116 KMLLID-YKENRLIACSLYQGIK--LLRLBDLKLGEYPYHKEHYLSGVNESG----- 167  
Db 118 MALLVDTYDDQLISCGSVNRGTQQRHVLPDSDSADIQSEVH---CMFSPSEESGQCPCD 174  
QY 168 --SVFGVIYSNLD--KLFIATVDGK--PEY-FTTISRKLTKNSEADFAFYVPHD 220  
Db 175 VVSALGAKVLLSEKDRFINFVGNTINSSYPGYSLSISVRLKETQ--DG-EKFLTDQ 231  
QY 221 EFVASMKIPIPSDFTIIPDP----DIYVYVGFSSGNFVFLTLOPEMVSPGSGTKEQV 276  
Db 232 SYI-----DVLPEFDSYPIKIHFASNNHFIYLTVOKE-----TLDQTF 273  
QY 277 TSKLVRCKEDTAFNSYVEVPIGC-----ERSGVE--YRLQAAVLSKAGAVLGRITLG 327

Db 274 HTRIRFCVSDSLHSEMPLECIILTEKRRKSTREEVFNILQAAVSKPGANLAKQIG 333  
Qy 328 VHPDDLLTVFSKGRKKSIDSALCIFILKQINDRIKERLQSCYRGEGTLDLAWLK 387  
Db 334 ASFSDILFGVFAQSPDSAEPAVSAVCAFFIKYNDFFNKIVNK----- 379  
Qy 388 VKDIPCSALLTDDNFCGLDMNAPL-----GVSMDVRGIPVFTEDDR 431  
Db 380 -NNVRLQHFYGNHEHC---FNRTLLNSSGCEARSDEVR--TEFTALQVLDLFWRL 433  
Qy 432 ----MTSVIAYVYKNSLAPVGTSGKSLKIRVDGRNALQVETVQVDPGPVLRDMAF 487  
Db 434 NQVLLTSIGTFIKGLDTIANLGTSEGRF-----MQVV----- 465  
Qy 488 SKDHEQIYMSERQLTRVPVSCGQYQSCGCEGLSGDPHCGVCVHLNCTRKERCERSKE 547  
Db 466 ----- 465  
Qy 548 PRFASSEMQCVRLTVPNNISVSQYNVLLVLETYNVPELSAGVNCFTFEDLSEMDGLVVG 607  
Db 466 -----LSTAHLTFRVN-----FLDSDHPVS----- 486  
Qy 608 NQIQYSPAKEVPRITENGDDHHVQLKSKETGTFASTSFVFNCSVHNSCLSCVE 667  
Db 487 -----f-----PEVIVEH----- 493  
Qy 668 SPYRCHWKVRHVCTHDPKCTSQEGRVKLPEDCPQLLRVDKILVPVEVIKPTLKAQNL 727  
Db 494 ----- 493  
Qy 728 PQQSGORGECILNIQSGEQRVPALRFNSSFVQCNQTSYSYEGMEINLPLVELTVVWNG 787  
Db 494 ---PSNQNGYTLVVT-----GKKITKIP--LNLGLCG 520  
Qy 788 HFNIDNPAQNKVHLKCGAMRESGLCKADDPFACWCGPGGCTLRQCHPQESOWLE 847  
Db 521 HF-----QSCQCLSAFYFQCGWCH--NOCVRFDECPs--GTWTQ 557  
Qy 848 LSGAKSKTNPRITEIPVTPREGGKVTIRGENLGER-----DIASHVKVAGVCS 902  
Db 558 -----EICLPVAVYKVPFTSAPLEGGLVTCGWDG--FRKNKFDLRTKVLGNESC 609  
Qy 903 PLVDGVIPIAEQIVCEMEAGPSHAGVEICVAVCRPEFMARSQLYFMTLTSLDKPS 962  
Db 610 TLTLSSTTNLAKTVGPM--SEFN--VSVIINSRE---TTQVSASFYDVPVITSIPR 664  
Qy 963 RGPMSGTQVITGTNLNAGSNVVMFGKPCPLFHRSPSYVNCV---TSSDEVLEMKV 1019  
Db 665 YGPAQAGTLLTLGKYLNSGNSRHISIGGKTCTLSKVSIDSILECYTPAQTSD---BFPV 721  
Qy 1020 SVQVDRAKIHQDLVQVYVEDPTVIRIEPWSIVSGNTPIAVWTHLDLQNPQ--IRAKH 1077  
Db 722 KLKIDILAN--RETSFSFREDPVVTEIHTPSFISGGSTIIGTILNSVSLFKVLVDVHE 780  
Qy 1078 GKKEHINICVLNATE--MTCQAPAL--ALGPDHQSGLTERPEEFGLDNVQSLILNKTN 1135  
Db 781 VGVNYTVACQHRNSSEIICCTPSLKQLG---LQLPLKTKAF--FLDDGILS-----KHFD 831  
Qy 1136 FTYYPNVFPAQSPGILELKPPTPIILKGNLIPPVAGGNVKNLYTVLVEKEPC----- 1190  
Db 832 LTYVHNVFPEFPKPVFMISGMNENNVKGNIDPEAVKGEV-----LKGVSQCSLHW 886  
Qy 1191 -----TVTVDVOLLCSNPLGRHKVMARVGMEYSPGMVYIAPDPSLPAIVSI 1244  
Db 887 HSGAVLCTVPSDDLKNSLSEINWQVSVTV-----LGKVIQVQDNFAGLIGAVSI 940  
Qy 1245 AGGLLIIFIVAVLIAYKRGRESDLTLKRLQW-----QMDNLESVALECKEAPAELO 1297  
Db 941 S-----VVLLLSGLFLWKRKRKDKLSELVYDARVHTPHLDRLVS-----ARSV 987  
Qy 1298 TDHIELSDLDGAGIPFLDRTYTMVLFPGIEHDPVLRDLVLP 1341

Db 988 SPTEMVSN-----ESVDYRATFPEDQFPNNSONGACRQVQYP 1025  
RESULT 4  
MET\_RAT  
ID MET\_RAT STANDARD; PRT; 1382 AA.  
AC P97523; P97579; Q63119; Q63964;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hepatocyte growth factor receptor precursor (EC 2.7.1.112) (Met proto-  
oncogene tyrosine kinase) (c-met) (HGF receptor) (HGF-SF receptor).  
GN MET.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
SEQUENCE FROM N.A.  
RP STRAIN=Sprague-Dawley; TISSUE=Liver;  
RC MEDLINE=97419268; PubMed=9271668;  
RX Wallenius V.R., Rawet H., Skritic S., Helou K., Qiu Y., Levan G.,  
RA Ekberg S., Carlsson B., Isaksson O.G.P., Nakamura T., Jansson J.-O.;  
RT "Chromosomal localization of rat hepatocyte growth factor (Hgf) and  
HGF receptor (Met) and characterization of HGF receptor cDNA.";  
RL Mamm. Genome 8:661-667(1997).  
RN [2]  
SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RP STRAIN=Sprague-Dawley; TISSUE=Kidney;  
RC MEDLINE=9706137; PubMed=8853431;  
RX Liu Y., Tolbert E.M., Sun A.M., Dworkin L.D.;  
RT "Primary structure of rat HGF receptor and induced expression in  
glomerular mesangial cells.";  
RL Am. J. Physiol. 271:F679-F688(1996).  
RN [3]  
SEQUENCE OF 364-495 FROM N.A.  
RP STRAIN=Sprague-Dawley; TISSUE=Gastric mucosa;  
RC MEDLINE=94220137; PubMed=8166728;  
RX Tsujii M., Kawano S., Tsuji S., Ito T., Hayashi N., Horimoto M.,  
RA Mita E., Nagano K., Masuda E., Hayashi N., Fusamoto H., Kamada T.;  
RT "Increased expression of c-met messenger RNA following acute gastric  
injury in rats.";  
RL Biochem. Biophys. Res. Commun. 200:536-541(1994).  
RN [4]  
SEQUENCE OF 851-1002 FROM N.A.  
RP TISSUE=Intestine;  
RC MEDLINE=95354758; PubMed=7628535;  
RX Pepper M.S., Soriano J.V., Menoud P.-A., Sappino A.-P., Orci L.,  
RA Montesano R.;  
RT "Modulation of hepatocyte growth factor and c-met in the rat mammary  
gland during pregnancy, lactation, and involution.";  
RL Exp. Cell Res. 219:204-210(1995).  
RN [5]  
SEQUENCE OF 1129-1267 FROM N.A.  
RP Kikuchi Y.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: RECEPTOR FOR HEPATOCYTE GROWTH FACTOR. HAS A TYROSINE-  
PROTEIN KINASE ACTIVITY.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
tyrosine phosphate.  
CC -!- SUBUNIT: HETERODIMER FORMED OF AN ALPHA CHAIN (50 kDa) AND A BETA  
CHAIN (145 kDa) WHICH ARE DISULFIDE LINKED (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, LIVER AND  
KIDNEY. ALSO EXPRESSED IN STOMACH, INTESTINE, SPLEEN, TESTIS AND  
BRAIN. NOT EXPRESSED IN HEART OR MUSCLE.  
CC -!- DEVELOPMENTAL STAGE: EXPRESSION IS DOWN-REGULATED DURING PREGNANCY  
AND IS VIRTUALLY UNDETECTABLE DURING LACTATION. EXPRESSION  
PROGRESSIVELY INCREASES POST-LACTATION.  
CC -!- INDUCTION: BY INTERLEUKIN-6 AND ACUTE ACID-INDUCED GASTRIC INJURY.  
CC -!- INHIBITED BY PROLACTIN.  
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
CC -!- SIMILARITY: CONTAINS 1 SEVA DOMAIN.





FT SIGNAL 1 24  
 FT CHAIN 25 1400  
 FT CHAIN 25 304  
 FT CHAIN 310 1400  
 FT DOMAIN 25 957  
 FT TRANSMEM 958 978  
 FT DOMAIN 979 1400  
 FT DOMAIN 58 507  
 FT DOMAIN 1082 1345  
 FT BIND 1088 1096  
 FT BINDING 1114 1114  
 FT ACT SITE 1208  
 FT MOD\_RES 1353 1353  
 FT MOD\_RES 1360 1360  
 FT CARBOHYD 66 66  
 FT CARBOHYD 419 419  
 FT CARBOHYD 458 458  
 FT CARBOHYD 488 488  
 FT CARBOHYD 654 654  
 FT CARBOHYD 720 720  
 FT CARBOHYD 841 841  
 FT CARBOHYD 897 897  
 FT VARSPLIC 884 932  
 FT VARIANT 322 322  
 FT SEQUENCE 1400 AA; 152227 MW; E8782D07E4D2568A CRC64;

Query Match 4.3%; Score 427; DB 1; Length 1400;  
 Best Local Similarity 19.7%; Pred. No. 2.9e-19;  
 Matches 329; Conservative 196; Mismatches 553; Indels 592; Gaps 73;

QY 38 RSFVTFRGPAGFNLVVDERTGTHLYLGAIVRIYKLSDDLKVLVTHETGPDENPKCYP 97  
 DB 56 QAMVTYEG-----DRNESAVFAIRNRLHVLGPDLSQVSLATGPAGD-FGC-- 101  
 QY 98 PRIVOTC-----NEPLTTNNVNKMLLDYKENRLIACGSLYQIGICKLLRLE---DLF 147  
 DB 102 ---QTCAACGPGPHGPGDGD--TKVLVDLPALPALVSCGSLQGRFCFLHLEPOGTAV 155  
 QY 148 KLGEF-----YHKKEHLYSGVNSGSGFVIVSVNSLDDKLFIAFADVG--KPEYFP-T 198  
 DB 156 HLAAPACPLFAHNRPDPCDCVAPSLGTRVTVEQQAQSYFYVASSLDAAVAGSFSRPS 215  
 QY 199 ISSRKLTKNSEADGMPAYVFHDEFVASMKIIPSDTFIIP-----DFDIYVYVGFSSGNFV 254  
 DB 216 VSIRL--KADASG-----FAPGFVA-----LSVLPHLVSYSEIYVHSHFTGAFV 259  
 QY 255 YFLTLOPEMVSPGSTTKQVYTSKLVRLCKEDTAFNSVVEVPICGE-----RS 303  
 DB 260 YFLTQVPSV-----TDDPSALHTRLARLSATEPELGDYRELVDLCRPAPKRRRGAPEG 314  
 QY 304 GVEYRLQAAVLSKAGAVLRTGLVHPDDLLFTFVSKQGRKMKSLDESALCIFILKQI 363  
 DB 315 QQYFVPLQVAHAPVGAQALATSLSAEQGVLFVGVFTGKGPGVGVGPNVWCAFPIDLL 374  
 QY 364 NDRIKERLOS CVRGETLDAWLKVKDIIPCSSAL-----LTDDNFC-----GLD-- 408  
 DB 375 DTLIDEGVERC-----CESVHPGLRRGLDFFQSPSFCFNPPLGLEAL 416  
 QY 409 -----MNAPLGVS-----DMVRGI--PVFTEDRDMTSVIAVYVKNHSLAPVGTGSKG 454  
 DB 417 SPNTSCRHFLVSSFSRVDLNLGLGPV-----QVTALYTRLDNVTVAHMGTDGR 470  
 QY 455 LKKIRVDGPRGNALQVETVQVDPGVPLRDMAFSKDHEQLYIMSERQLTRVPVESCQVQ 514  
 DB 471 I-----LQVELVR----- 478  
 QY 515 SCGECGLSGDPHCGWCWVLTNCTRKRCERSKEPRFASEMKQCQVLTVHPNNISVQYN 574  
 DB 479 -----SLN 481

QY 575 VLLVLETYNVPELSAGVNCVTFEDLSEMDGLVVGNIQICYSAPAAKEVPRIITENGDDHHVQV 634  
 DB 482 YLLVSNFSLGDSG-----QPVQDVSL-----GDH----- 508  
 QY 635 LQLKSETGMTFASTSVFVYVNCVSHNSCLSCVBSPYRCHWKYRHVCTHDPKTCFQEGR 694  
 DB 509 -----LLFASGDQVF-----QVPIRPGC--RHFLT----- 532  
 QY 695 VKLPEDCPQLLRVDKILVPVEVIKPTLKAKNLPQPSQGRGYECILNIQSGEQRVPAIR 754  
 DB 533 ----- 532  
 QY 755 FNSSSVQCNTSYSEGMENNLPELVTVVMNGHFNIDNPAQNKVHLKCGAMRSCGLC 814  
 DB 533 -----CGR 536  
 QY 815 LKADDPFACWCGPGGCTLRHCPCAOESOWLELSGAKSKCTNPRITIIIPVGTREGGT 874  
 DB 537 LRAWHFMGCCWC--GNMCGQQKECPG-----SWQODHCP-PKLTFFHPSGSLRGST 585  
 QY 875 KVTIRGENLGFPRDI---ASHVKVAGVCSPLVDGY-----IP---AQIVCEMGEA 921  
 DB 586 RLTLGSGNFYLPSPGLVPEGTHQVTVGQSPCRPLPKDSSKULPRPKDFVEFECEL-EP 644  
 QY 922 KPSQHAGFVBIQVAVCR---PEFMARSSQL---YFMTLTLSLKPGRGMSGGTQVTI 974  
 DB 645 LGTQAVGPTVSLVTVNMPGKHPRVDGTVLRGFSFMEPVLLIAVQLPFGPRAGGTCLTL 704  
 QY 975 TGNLNAGSNVVMFGKQPCLFHRRSPSYIVCNTTSSDEVLEMKVSVQVDRAKIHODLVF 1034  
 DB 705 EGQSLSVGTSAVLVNGTECLLARVSEGLLCATPGATVASVPLSLQVGAQVPGSWTF 764  
 QY 1035 QYVEDPTVIBBEMSVSGNTPRIAVNGTHLDLQNPQIRAKHGGKEHINICEVLNATEM 1094  
 DB 765 QYREDPVVLSISNCGYI--NSHITICQHLTSAWH-LVLSFHDG-----LRAVES 812  
 QY 1095 TCQAPALALGPHQSDLTERPEERGFILDNVQSLILNKT-----NFTYYPNP 1142  
 DB 813 RCERQL---PEQ--LCRLPE---YVVRDPQGVVAGNLSARGDGAAGFTPLPGFRLPPP 863  
 QY 1143 VFRAGPS-GILELKPOTPII---LKGKNLIPPVAGNVKLNLYTVLGEKPCCTVTVSDVQ 1198  
 DB 864 ---HPSANLVLPLKEHAIKFEYI GLGAVADCVGIN-----VTVGESQCHFERGDM 913  
 QY 1199 LICESPNLI-----GRHKVMARVCGMEYSPGMVYIAPDS-PLSLPAIVSTA 1243  
 DB 914 VVCPPLPSLQGDGAPLQVCDGEGCHILGRV---VRPG-----PDGVPOSTLLGILLP 964  
 QY 1244 VAGGLLIIFTVAVLIIAYKRKSRESDLTKRLQMQMDNLESRALECKEAFELQTDIHEL 1303  
 DB 965 L---LLIIVAAALATVFSYWWRRKQLVLP-----PNLNDLASLD-QTAGA---TPLPL 1011  
 QY 1304 TSOLD---GAGIFP---LDYRTYTMVLPFGIEHPVLRLDLEVPGYORVEKGLKLFQA 1357  
 DB 1012 YSGSDYRSGLALPAIDGLDSTTCVHGSAFSDSEDESC-----VPLLRKESQLRLDSAL 1066  
 QY 1358 LINNKVPLLSFIRTL-ESOR-----SFSMRDRGNVASLIMTVLQSKLEYATDV--LKOLL 1409  
 DB 1067 LAEVKDVLIIPHERVVTSHSDRVIGKHGFGVYHGEYIDQAQNRIOCAIKSLSRITEMQQVE 1126  
 QY 1410 ADLIDKML-ESKNHPKLL-----LRRTESVAEKMLTNMFTFLLYKFLKECAGEPLFSLFC 1463  
 DB 1127 AFLREGLLMGLNHPNVLLALIGLIMLPPEGLPHVLLPYMCHGDLLOFIRSQRNPVTKDLI 1186  
 QY 1464 AIKQOMEKGPIDAITGEARYSLSEDKLIROQIDYKTLVLSCVSPDNANSPEVPVKILNCD 1523  
 DB 1187 SFGQLVARG-----MEY-LAEQKFVHRDLAARNCM-----DESFTVKVADF 1228  
 QY 1524 TITQVKEKILDAIFKNVPCSHRPAKADMDLEWROGSGARMILQDEDTTK 1573  
 DB 1229 ---LARDILDRYYSVQ-QHR---HARLPVKWM-----ALESQTVRFVTTK 1267

## RESULT 6

ID - RON\_MOUSE STANDARD; PRT; 1378 AA.  
 AC Q62190; Q62555;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Macrophage-stimulating protein receptor precursor (EC 2.7.1.112)  
 DE (MSP receptor) (p185-Ron) (Stem cell-derived tyrosine kinase).  
 GN MSTIR OR RON OR STK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94250897; PubMed=8193352;  
 RA Iwama A., Okano A., Sudo T., Matsuda Y., Suda T.;  
 RT "Molecular cloning of a novel receptor tyrosine kinase gene, STK,  
 RT derived from enriched hematopoietic stem cells.";  
 RL Blood 83:3160-3169(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ;  
 RX MEDLINE=98127434; PubMed=9467940;  
 RA Waltz S.E., Toms C.L.V., McDowell S.A., Clay L.A., Muraoka R.S.,  
 RA Air E.L., Sun W.Y., Thomas M.B., Degen S.J.F.;  
 RT "Characterization of the mouse Ron/Stk receptor tyrosine kinase  
 RT gene.";  
 RL Oncogene 16:27-42(1998).  
 CC -!- FUNCTION: RECEPTOR FOR MACROPHAGE STIMULATING PROTEIN (MSP). HAS A  
 CC TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBUNIT: HETERODIMER FORMED OF AN ALPHA CHAIN AND A BETA  
 CC CHAIN WHICH ARE DISULFIDE LINKED (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- PTM: THE TWO SUBUNITS ARE PROTEOLYTICALLY DERIVED (BY SIMILARITY).  
 CC -!- PTM: PHOSPHORYLATED IN RESPONSE TO LIGAND BINDING (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 1 SEMA DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X74736; CAA52754.1; -;  
 CC EMBL; U65949; AAC39953.1; -;  
 CC HSP; P11362; IFGK.  
 CC MGD; MGI:99614; Mstir.  
 CC InterPro; IPR000719; Euk\_pkinase.  
 CC InterPro; IPR002909; IPT\_TIG.  
 CC InterPro; IPR002165; Plexin\_repeat.  
 CC InterPro; IPR001627; Sema.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC InterPro; IPR001245; Tyr\_pkinase.  
 CC Pfam; PF00069; pkinase; 1.  
 CC Pfam; PF01403; Sema; 1.  
 CC Pfam; PF01437; PSI; 1.  
 CC Pfam; PF01833; TIG; 3.  
 CC ProDom; PD000001; Euk\_pkinase; 1.  
 CC SMART; SM00220; S\_TKC; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;  
 CC Glycoprotein; Signal; Phosphorylation; ATP-binding.  
 CC SIGNAL 1 23  
 CC CHAIN 24 1378 MACROPHAGE-STIMULATING PROTEIN RECEPTOR.

CHAIN	25	305	MACROPHAGE-STIMULATING PROTEIN RECEPTOR, ALPHA CHAIN (POTENTIAL).
FT	311	1378	MACROPHAGE-STIMULATING PROTEIN RECEPTOR, BETA CHAIN (POTENTIAL).
FT	25	960	EXTRACELLULAR (POTENTIAL).
FT	961	981	POTENTIAL.
FT	982	1378	SEMA.
FT	60	509	CYTOPLASMIC (POTENTIAL).
FT	1059	1322	PROTEIN KINASE.
FT	1065	1073	ATP (BY SIMILARITY).
FT	1091	1091	ATP (BY SIMILARITY).
FT	1185	1185	BY SIMILARITY.
FT	1330	1330	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	1337	1337	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	91	91	N-LINKED (GLCNAC) (POTENTIAL).
FT	391	391	N-LINKED (GLCNAC) (POTENTIAL).
FT	460	460	N-LINKED (GLCNAC) (POTENTIAL).
FT	490	490	N-LINKED (GLCNAC) (POTENTIAL).
FT	656	656	N-LINKED (GLCNAC) (POTENTIAL).
FT	722	722	N-LINKED (GLCNAC) (POTENTIAL).
FT	845	845	N-LINKED (GLCNAC) (POTENTIAL).
FT	901	901	N-LINKED (GLCNAC) (POTENTIAL).
FT	713	713	VG -> IA (IN REF. 1).
FT	719	719	V -> A (IN REF. 1).
SQ	SEQUENCE	1378 AA; 150538 MW; FCSF932377BS7009 CRC64;	

  

Query Match	4.0%	Score 398.5;	DB 1;	Length 1378;
Best Local Similarity	19.5%	Pred. No. 2e-17;		
Matches 271;	Conservative 164;	Mismatches 436;	Indels 517;	Gaps 57;

  

QY	57	DERTGHYLVGAVNRIYKLSLKVLTHTGPDENPKYPRIVQTC-----NEPLT 109
DB <td>65 <td>DSTNSAVFVATRNHLVLPDLQFTENLTGPG-IGNPGC-----QTASCAGPGHPGPK 117</td> </td>	65 <td>DSTNSAVFVATRNHLVLPDLQFTENLTGPG-IGNPGC-----QTASCAGPGHPGPK 117</td>	DSTNSAVFVATRNHLVLPDLQFTENLTGPG-IGNPGC-----QTASCAGPGHPGPK 117
QY <td>110 <td>TNNVNMMLLDYKENRLIACSLYQGIKLLRLDLFKLGPYKKEHY-----LSGVN 164</td> </td>	110 <td>TNNVNMMLLDYKENRLIACSLYQGIKLLRLDLFKLGPYKKEHY-----LSGVN 164</td>	TNNVNMMLLDYKENRLIACSLYQGIKLLRLDLFKLGPYKKEHY-----LSGVN 164
DB <td>118 <td>DTDTL--VLVMEPGLPALVSCGSLQGRCLHEL-----EPKGLALHAPACLFSA 168</td> </td>	118 <td>DTDTL--VLVMEPGLPALVSCGSLQGRCLHEL-----EPKGLALHAPACLFSA 168</td>	DTDTL--VLVMEPGLPALVSCGSLQGRCLHEL-----EPKGLALHAPACLFSA 168
QY <td>165 <td>ES-----GSVFG--VIVSYSLNLDKLFATAVDGKPEYFPTISSKLTKNSEADGM 213</td> </td>	165 <td>ES-----GSVFG--VIVSYSLNLDKLFATAVDGKPEYFPTISSKLTKNSEADGM 213</td>	ES-----GSVFG--VIVSYSLNLDKLFATAVDGKPEYFPTISSKLTKNSEADGM 213
DB <td>169 <td>NKPEACTDCVASPLGTRVTVVVEQGHASVYVASSLD--PELAASFSPRSV----- 217</td> </td>	169 <td>NKPEACTDCVASPLGTRVTVVVEQGHASVYVASSLD--PELAASFSPRSV----- 217</td>	NKPEACTDCVASPLGTRVTVVVEQGHASVYVASSLD--PELAASFSPRSV----- 217
QY <td>214 <td>FAYVHFDEFVASMIKIPSDTTIIPDFD-----IYVYVGFSSGNFVFELTQ 260</td> </td>	214 <td>FAYVHFDEFVASMIKIPSDTTIIPDFD-----IYVYVGFSSGNFVFELTQ 260</td>	FAYVHFDEFVASMIKIPSDTTIIPDFD-----IYVYVGFSSGNFVFELTQ 260
DB <td>218 <td>-----SIRRLKSDTSGFGFPFSLSVLPKYLASYLYKYVYFSGDFVFYELTVQ 266</td> </td>	218 <td>-----SIRRLKSDTSGFGFPFSLSVLPKYLASYLYKYVYFSGDFVFYELTVQ 266</td>	-----SIRRLKSDTSGFGFPFSLSVLPKYLASYLYKYVYFSGDFVFYELTVQ 266
QY <td>261 <td>PEWSPPGSTTKEQVYTSKLVRLCKEDTAFNSYVEVPIGC-----ERSGV-----EYRL 309</td> </td>	261 <td>PEWSPPGSTTKEQVYTSKLVRLCKEDTAFNSYVEVPIGC-----ERSGV-----EYRL 309</td>	PEWSPPGSTTKEQVYTSKLVRLCKEDTAFNSYVEVPIGC-----ERSGV-----EYRL 309
DB <td>267 <td>PISVTSPFSAH-----TLVRLNAVEPIGDYRELVLVDCHFAPKERRRGAPEGTQYPV 321</td> </td>	267 <td>PISVTSPFSAH-----TLVRLNAVEPIGDYRELVLVDCHFAPKERRRGAPEGTQYPV 321</td>	PISVTSPFSAH-----TLVRLNAVEPIGDYRELVLVDCHFAPKERRRGAPEGTQYPV 321
QY <td>310 <td>LQAAYLKAGAVLGRGLVHPDDLLFTVFSKGQRKMKSLDESALCIFIILKQINDRIKE 369</td> </td>	310 <td>LQAAYLKAGAVLGRGLVHPDDLLFTVFSKGQRKMKSLDESALCIFIILKQINDRIKE 369</td>	LQAAYLKAGAVLGRGLVHPDDLLFTVFSKGQRKMKSLDESALCIFIILKQINDRIKE 369
DB <td>322 <td>LQAAHAPVDAKLAVELSISEQEVLFGVTVTKDGGSGMSPNSVVCAPFIYHLNILEE 381</td> </td>	322 <td>LQAAHAPVDAKLAVELSISEQEVLFGVTVTKDGGSGMSPNSVVCAPFIYHLNILEE 381</td>	LQAAHAPVDAKLAVELSISEQEVLFGVTVTKDGGSGMSPNSVVCAPFIYHLNILEE 381
QY <td>370 <td>RLQCYRGEGLDLAWLKVDPICSSALLTIDNFCGLDMNAPLGVSVMVRGIPVETEDR 429</td> </td>	370 <td>RLQCYRGEGLDLAWLKVDPICSSALLTIDNFCGLDMNAPLGVSVMVRGIPVETEDR 429</td>	RLQCYRGEGLDLAWLKVDPICSSALLTIDNFCGLDMNAPLGVSVMVRGIPVETEDR 429
DB <td>382 <td>GVEYCHSSN-----SSSLLS-----RGLDFP----- 403</td> </td>	382 <td>GVEYCHSSN-----SSSLLS-----RGLDFP----- 403</td>	GVEYCHSSN-----SSSLLS-----RGLDFP----- 403
QY <td>430 <td>DRMTSVIAYVYVKNHSLAFVGTGSKGLKKIRVDGPRGNALQYETVQVYVDFGVLDRMAFSX 489</td> </td>	430 <td>DRMTSVIAYVYVKNHSLAFVGTGSKGLKKIRVDGPRGNALQYETVQVYVDFGVLDRMAFSX 489</td>	DRMTSVIAYVYVKNHSLAFVGTGSKGLKKIRVDGPRGNALQYETVQVYVDFGVLDRMAFSX 489
DB <td>404 <td>-----QTSPFCNPFG----- 414</td> </td>	404 <td>-----QTSPFCNPFG----- 414</td>	-----QTSPFCNPFG----- 414
QY <td>490 <td>DHEQLYIMSERQLTRVPVESCQYOSCGELSGDGHGCVLHNTCTKRCERSKEPR 549</td> </td>	490 <td>DHEQLYIMSERQLTRVPVESCQYOSCGELSGDGHGCVLHNTCTKRCERSKEPR 549</td>	DHEQLYIMSERQLTRVPVESCQYOSCGELSGDGHGCVLHNTCTKRCERSKEPR 549
DB <td>415 <td>-----GEASGP-----SSRC----- 424</td> </td>	415 <td>-----GEASGP-----SSRC----- 424</td>	-----GEASGP-----SSRC----- 424
QY <td>550 <td>RFASEMKQCVRLTVHPNNISVSQVNVLLVLETVNVFELSAGVNCCTPEDLSEMDGLVVGNO 609</td> </td>	550 <td>RFASEMKQCVRLTVHPNNISVSQVNVLLVLETVNVFELSAGVNCCTPEDLSEMDGLVVGNO 609</td>	RFASEMKQCVRLTVHPNNISVSQVNVLLVLETVNVFELSAGVNCCTPEDLSEMDGLVVGNO 609
DB <td>425 <td>-----HYFPLMVRASFTRVDLFGLL-----GSVKVTAHVLTRIGNVTVV-- 464</td> </td>	425 <td>-----HYFPLMVRASFTRVDLFGLL-----GSVKVTAHVLTRIGNVTVV-- 464</td>	-----HYFPLMVRASFTRVDLFGLL-----GSVKVTAHVLTRIGNVTVV-- 464
QY <td>610 <td>IQCYSPAAREPRIITENGDDHHVQLKSKETGMTFASTSFVYVNCVSHNSCLSCVESP 669</td> </td>	610 <td>IQCYSPAAREPRIITENGDDHHVQLKSKETGMTFASTSFVYVNCVSHNSCLSCVESP 669</td>	IQCYSPAAREPRIITENGDDHHVQLKSKETGMTFASTSFVYVNCVSHNSCLSCVESP 669
DB <td>465 <td>----- 464</td> </td>	465 <td>----- 464</td>	----- 464





```
QY 9 TLLSHLLM---VGMGSST---LLTRQAPL--SOKORSFV---TPRGPAGGKFNHLVVD 57
D 17 TLVSHLSSQDIASSESSEQOCTREHPVIAFEDLKPWFVNFYYPG--VRDFSQALD 74
QY 58 ERTGHIYLGAVNRIYKLS--SDLKVLVTHETGPDENPKCYPPIVQTCNEPLITTTNNVK 116
D 75 PSRNOLIVGGRNLFRLSLANVSLLOATEWASSED-----RRSCQSGKTBEECQNYVR 129
QY 117 MLLIDYKENRLIACG-SLYQGIKCLRLLEDLKLGE-----PVHKHEHYLSGVNESG 167
D 130 VLIVSGR--KVPWCGTNAFSPVCSRQVGNLRTTEKINGVARCPDPRHNSTAVISSQ 187
QY 168 SVF-GVIVSYSLNDKLFATADVGKPEYFPPTISSRKLTKNSEADGMFAYVHDEFAVM 226
D 188 ELYAATVIDFSGRDAIYRSLG-SGPPLTAQYNSKWLNE-----226
QY 227 IKIPSDTTIIPDFIYVYVYGFSSGNFVFLTLQPEMVSPPGSTTKBQVYVTKLVLCKE 286
D 227 -----PNF-----VAAFIDIGLFAFFFLRENAVEHDCGRT---VY-SRVARVCKN 266
QY 287 DT-----AFNSYVEVPIGICRSQ---VEYRLLOAAVLSKAGVLGRTLGVH-PDDDL 334
D 267 DVGGFLEDTWTTFMKARLNCRSGEVFPFYNELQSAP-----HLPEQDL 312
QY 335 LFTVPSKGOKRWKSLDESALCIFILKQINDRIKERLQSCYRGEGLDLDLAWLKVKD-IPC 393
D 313 IYGVFT---TNVNSIAASAVCAFNLISLS-----KAFNGFPFYQENPRAALPIANPIP- 363
QY 394 SSALLTIDNF-CGL-----DNNAFLGVSDMVRGI---PVFTEDRDMT 433
D 364 -----NFQCGTLPETGPNENLTERSLQDAQRFLMSEAVQVTPPEPCVTQDSVRS 414
QY 434 SVIAVYKN-----HSLAFVGTSGKLKTRVDGPRG-NALQVETVOVDPGPV--LRDMA 486
D 415 HLVDVLQAKDLYHLVYLTGTSGLTKALSTASRLGRCYLEELHVLPPGRLEPLRSR 474
QY 487 FSKDEHQIYIMSEROLTRVPVPSGQYQSCGCEGLSGDPHCGV-----529
D 475 ILHSARALFVGLSDRVLRIPLERCSNAYSQGLGARDPYCGWDGKRLQCLSTLEDSSNMS 534
QY 530 -----CVLHN-----TCTRKERCERSEKPRRFASEM 556
D 535 LMIQNTTTPVRNVTDRDGGFPWSPKCEHLDDGNSGCLCRARCDSPRRCGGLG--592
QY 557 QCVRLTVHPNNLSVQYVLLVLETYNVPSLSAGVNTCTFEDISEMDGLVVGNOIQCYSPA 616
D 593 -CLGPSIHTAN--CSRNGAWTAWSSWAQCSSTSGI-----GFQV-RQRSCSNFA 637
QY 617 AKEVPRIITENGDDHVVQLQSKETGWTFASTF---VFY-----NCSVHNSCLSCV 666
D 638 PRHGRICVG-----KSRERFCNENTPCVPPIFNASGWSKCS--NNGCGGV 684
QY 667 ESPYRCHWCKYRHVCTH---DPKTCFQBRGVKLPCDPQLLRVDKIL--VPVEVIX---718
D 685 QSRRS--CENGSCPCGCVBPCTCN-----PEACPEVRNTPWLPVNVVQGA 734
QY 719 -----PITLAKNLP--OSQORGYE-----738
D 735 ROQRFRFTCRAP-LPDPHGLQFGKERTTCTCPADGTGACDTDALVEDLLRSGSTSPHT 793
QY 739 -----CILNTQGSQRVPALRFNSSVQOQNTSYSEGMENNLIPV--779
D 794 LMGHWATGPNSSCRSDCLGFRVVRKTRCTNPEPRNGGLPCVGDAAEYQDCNPQACPVRG 853
QY 780 --ELTVVMNGHFNIDNPAQNKVHLKCGAMRESG-----LCLKADP-----DFAAGW 825
D 854 AWSCTAWS-----QCSA--SCGGHYQRTSCTSPAPSPGEDIICLGL 894
QY 826 COGPGQCTLRHCPAQESQWLEL-----SGAKSK-----CTN 857
D 895 HTEALCS-TOACPEGWSLWNGVCTEDGAQSRSRSCBELLPGPGACVGNSSQSRPCFY 953
```

858 PRTEIIPVTPGREGTKVIRGENGLGFERDIASHVKVAGVECSPLVDGYIP-AEQIVC 916  
954 SEIPVILPASSVEE---TTSCGGFNL-----IHLIVTGVSCF-LVSGLLTLAVYLS 1001  
917 EMGEAKPSQAGVFEICVAVCRPEFMRSSQLY---FMTLTLSDLKP 961  
1002 QHCQ-RSQSESTLVH-----PATPNHLHYKGGTTPRNEKYTPMEFXTLKNKNLIP 1050

RESULT 9  
SM4B\_MOUSE STANDARD; PRT; 782 AA.  
AC Q62179;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).  
GN SEMA4B OR SEMAC OR SEMC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NMRI; TISSUE=Brain;  
RX MEDLINE=95267431; PubMed=7748561;  
RA Puschel A.W., Adams R.H., Betz H.;  
RT "Murine semaphorin D/collapsin is a member of a diverse gene family  
and creates domains inhibitory for axonal extension.";  
RL Neuron 14:941-948(1995).  
RN [2]  
RP INTERACTION WITH GIPC.  
RX MEDLINE=99253973; PubMed=10318831;  
RA Wang L.-H., Kalb R.G., Strittmatter S.M.;  
RT "A PDZ protein regulates the distribution of the transmembrane  
semaphorin, M-Semf.";  
RL J. Biol. Chem. 274:14137-14146(1999).  
CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO  
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.  
CC -1- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.  
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW  
LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH  
MODERATE LEVELS FROM THEN UNTIL BIRTH.  
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL: X85992; CAA59984.1;  
DR MGD; MGI:107559; Sema4b.  
DR InterPro; IPR003659; Plexin-like.  
DR InterPro; IPR002165; Plexin\_repeat.  
DR InterPro; IPR001627; Sema.  
DR Pfam; PF01403; Sema; 1.  
DR Pfam; PF01437; PSI; 1.  
DR SMART; SM00423; PSI; 1.  
DR Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;  
KW Developmental protein; Glycoprotein.  
FT NON TER 1 662 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN <1 662 POTENTIAL.  
FT TRANSMEM 663 683 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 684 782 SEMA.  
FT DOMAIN 202 492 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 548 608 PRO-RICH.  
FT DOMAIN 703 726

```
FT DISULFID 555 601 BY SIMILARITY.
FT CARBOHYD 12 12 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 15 15 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 526 526 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 574 574 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 639 639 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 782 AA; 86823 MW; 627A81FC8F8F7AC8 CRC64;

Query Match 3.1%; Score 308.5; DB 1; Length 782;
Best Local Similarity 20.9%; Pred. No. 5.3e-12;
Matches 156; Conservative 121; Mismatches 292; Indels 177; Gaps 32;

QY 36 KQSFVTFGEPAEGNHLVVDERTGHIYLGAVNRIYKLSDDLKVLVTHE-----TQDPE 90
DQ 1 EERLIRKFEAENISNTALLSQGKTLVYGAREALFALNSLFLPGGYQELLWSADA 60
QY 91 DNPCKYPPRIQVTCN---EPLTTNNVNMWLLIDYKENELIAGC-SLYQGICKLLRLLED 145
DQ 61 DRK-----QOCSPGKDPKDCQNYIKILL-PLNSHLLTCGTAAPFLCAYIHAS 111
QY 146 LFKLGPYHKHYLSGVNBSGSVFGVIVSYNLD-----DKLFIATA--VDGKPEY 195
DQ 112 -FTLAQ-----DEAGNVI-----LEDKGHCPEDPNFKSTALVVDG--EL 148
QY 196 F-----PTISRKLTKNSEADCMFAYVPHDEFVAMIKIPSDTTIIPDPDIY 244
DQ 149 YTCVTSFQNDPAISRQSSRPTKTESSLNWLQDPAFVASATSPESLGPSPGDDDKIYF 208
QY 245 VYGFSSGNFYVFLTLQPEMVPSPGTTKEQVYTSKLVRLCKETA-----FNSYVEV 296
DQ 209 PFSETQGEFF-----ENTIVSRVARVCKGDEGGRVLQQRWTSFLKA 252
QY 297 PIGCER--SGVEYRLLOAAYLSKAGAVLGRITLGVHPDD---DLLTFVSKGQKRKMSLD 351
DQ 253 QLLCSRPDGFFPNVLQDVE-----TLNPNQDWRKTLISGVFTSQWHR--GTTE 300
QY 352 ESALCIFILQINDRIKERLOSVCYRGEGTLDLAW-LVKVDIIPCSSALLTIDDNFCGLDMN 410
DQ 301 GSAICVFTM---ND-VOKAFDGLYKKNRRTQOQYTTETHQVTPRPGACITNSAREKIN 356
QY 411 APLGVSD-----MVRGIPVFTEDRDMTSLVIAI-----VYKNHSLAFVGTKS 452
DQ 357 SSLQLPDRVLNFKDHFMDQVRSRLLQLQPRARYQVAVHRVPLGHSYDVLFLGTGD 416
QY 453 GKLIKIRVDGPRGNALQYTVQVVDPGFVLDRMAFSDKHQOLYIMSERQLTRVPVESCGQ 512
DQ 417 GRHLKAVTLSSRVHII--EELQIPPOGQPVQNLILLDSHGGLLYASSHSVGVQVAVNCSL 474
QY 513 YQSCGEGCLSGDHPCHGW---CVLHNT-----CTRKECERSKEPREFA 552
DQ 475 YPTCGDCLLARDPYCAWTGSACRLASLYQDPLASRPWTQDIEGASVRELKNSYKARFL 534
QY 553 SEMKQCVRLTVHPNNISVQVNVLLVLET-----YXNPVLSAGVNCCTPEDSEMGLVVGN 608
DQ 535 VPGKPKQVOIQPNTVNTLACPLLSNLATRLWVINGAPVNASASCRV--LPTGDLILLVGS 592
QY 609 Q-----IQCVS-----PAAKEVPRIITENGDDHHVQVLQKSKETGTFASTSFVYNC 656
DQ 593 QQGLGVPCQWSIEBGFQQLVASCYPEVMEEG-----VMDQKNQDGTGTVIINTSRVSAPA 647
QY 657 SVHNSCLSCVESPYRCHCKYRHVCT 682
DQ 648 GGRDSW-----GADKSYWNEFLVMCT 668

RESULT 10
SMSA_HUMAN
ID SMSA_HUMAN STANDARD; PRT; 1074 AA.
AC Q13591; Q60408;
```

```
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 5A precursor (Semaphorin F) (Sema F).
GN SEMA5A OR SEMAF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9812554; PubMed=9464278;
RA Simmons A.D., Puschel A.W., McPherson J.D., Overhauser J., Lovett M.;
RT "Molecular cloning and mapping of human semaphorin F from the Cri-du-
RL Biochem. Biophys. Res. Commun. 242:685-691(1998).
RN [2]
RP SEQUENCE OF 1-494 FROM N.A.
RA Kalicki J., Harmon G.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY ACT AS POSITIVE AXONAL GUIDANCE CUES.
CC -!- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 TSP TYPE-1 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U52840; AAC09473.1; -.
DR EMBL; AC004615; AAC14668.1; -.
DR Genew; HGNC:10736; SEMA5A.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 6.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF01437; PSI; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00423; PSI; 1.
DR PROSITE; PS50092; TSP1; 5.
DR Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 1074 SEMAPHORIN 5A.
FT DOMAIN 23 968 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 969 989 POTENTIAL.
FT DOMAIN 990 1074 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 226 507 SEMA.
FT DOMAIN 540 609 TSP TYPE-1 1.
FT DOMAIN 595 652 TSP TYPE-1 2.
FT DOMAIN 653 704 TSP TYPE-1 3.
FT DOMAIN 707 765 TSP TYPE-1 4.
FT DOMAIN 784 840 TSP TYPE-1 5.
FT DOMAIN 841 897 TSP TYPE-1 6.
FT DOMAIN 898 952 TSP TYPE-1 7.
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 933 933 N-LINKED (GLCNAC. .) (POTENTIAL).
```





QY 210 ADGMEYVHDFVASMIPKIPSDTETIIPDFIYVYVGFSSGNFVYFLTLQEMVSPPGS 269  
Db 221 -----HD---SKWFKP-----YFVHAVEMGSHVYFFREIAM-----BF 252  
QY 270 TTKEQVYTSKLVRLCKEDTA-----FNSYVEVPIGCSRG-----VEYRLLOAAYLSK 317  
Db 253 NYLEKVVSVSRVAVCKNDVGGSPRVLEKQWTSFLKARLNCSPVGDHSHFVFNLOAV----- 308  
QY 318 AGAVLGRITGLVHPDDLLFTVFSKQKRWKSLDESALCIFILKQINDRIKRLQSCYRG 377  
Db 309 TGVV---SLGGRP---VLAVERF---TPSNIPGSAVCAFDLTQVAAVFEGR-----FRE 354  
QY 378 EGTLDLAWLKVD--IP-----CSSALLTIDNFCGLDMA-----PLGVSDMVVRGIPVFT 426  
Db 355 QKSPESITVPVEDQVPRPRGCCAA-----PCMQYNASSALPDDILNFVKTHPLMD 406  
QY 427 ED-----RDRMTSVIAYV---YKNSHLAFVGTGSKLKKIRV-----D 461  
Db 407 EAVPSLGHAPWILRTLMRHQLTRVAVDVAGPWGNTQVTVFLGSEAGTVLKLVRPNASTS 466  
QY 462 GPRGNALQVETVQVDDP-----GVLRDMAFSKDHQLYIMSERQLTRVPVESC 510  
Db 467 GTSGLSVLEEFETTRPRDRCRPGGGETQRLSLLELDAASGGLLAAPRCVVRVPEVARC 526  
QY 511 GOYQSC-GECLSGDPHCGW-----CVLHNTCTR 538  
Db 527 QOYSGCMKNCISQDPYCGWAPDGGCIFLSPGTR 560  
RESULT 13  
SM4A\_HUMAN STANDARD; PRT; 762 AA.  
ID SM4A\_HUMAN  
AC Q9H3S1;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Semaphorin 4A precursor (Semaphorin B) (Sema B).  
GN SEMA4A OR SEMB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,  
RA Saito T.;  
RT "Human semaphorin B";  
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO  
CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 SEMA DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; AB029394; BAB20087.1; .  
CC InterPro; IPR002165; Plexin\_repeat.  
CC InterPro; IPR003659; PSI.  
CC Pfam; PF01437; Plexin\_repeat; 1.  
CC SMART; SM00423; PSI; 1.  
KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;  
KW Neurogenesis; Developmental protein; Glycoprotein.  
FT SIGNAL 1 32  
POTENTIAL.

FT CHAIN 33 762 SEMAPHORIN 4A.  
FT DOMAIN 33 684 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 685 705 POTENTIAL.  
FT DOMAIN 706 762 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 240 518 SEMA.  
FT DOMAIN 574 632 IG-LIKE C2-TYPE DOMAIN.  
FT DISULFID 581 625 BY SIMILARITY.  
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 762 AA; 83195 MW; 6A738E4BEB83D4 CRC64;  
Query Match 2.6%; Score 260.5; DB 1; Length 762;  
Best Local Similarity 21.7%; Pred. No. 6.5e-09;  
Matches 158; Conservative 105; Mismatches 283; Indels 181; Gaps 33;  
QY 5 PWNWTCI-----LSHLLM-----VGMGSSTLTITROPAPISQKORSFVTRGRGPAEGFNHLV 56  
Db 10 PWSLLGLFLQLLQLLPTTTAGGGGQGPMPRVRYVYAGDERRALLSFHQKGLQDFDTLL 69  
QY 57 DERTGHYLGAVNRVYKLSDD-----LKVLTHTETGPDENPKYPRIVQTCNEPLTT 110  
Db 70 SGGNTLYVGAREAILALDIDQDPGVPRLKMIWPAS-DRKKSEC---AFKKKNE----- 121  
QY 111 TNNVNM-LIIDIYKENRLIACGSL-YQICIKLRLLEDLFKLK-----EPY 153  
Db 122 TQCFNFRVLVSINVTHLYTCGTFAPSPACTFTELQDSYLLPISDKVMEKGQSPDP 181  
QY 154 HKKEHYL-SGVNBSGVFVIVSYNLDLFTATAVDGKPEYPTTSSRLTKNSEADG 212  
Db 182 EKHTAVLDGMLYSGTWNFLGS-----EPILMRTLGSQPVLK---TDN 222  
QY 213 MFAYVFDH-BFVASMIKIPSDTETIIPDFIYVYVGFSSGNFVYFLTLQEMVSPPGSTT 271  
Db 223 FLRWLHHDASFVAA---IPSTQV-----VYFFETASEFDEF----- 257  
QY 272 KEQVYTSKLVRLCKEDTAFTNSYVEVPIGCR-----SGVEYRLLOAAYLSKAGAVLGR 324  
Db 258 -ERLHTRVARVCKND-----VGKLLQKKWTFLLKAQLLSAPSGSCPSTSSA 306  
QY 325 TLGVHPD--DDLFTVFSKQKRWKSLDESALCIFILKQINDRIKRLQ-----S 373  
Db 307 TRSCSPPILPQLPSTQSSPPSGVGGTRSSAVCAFSILLDIERVFKGFKELNKETSRWT 366  
QY 374 CYEGEGTLDLAWLKVDIPCS-----SALLTIDNFCGLDMAPLGVSDMVVRGIPVETE 427  
Db 367 TYRGPEP-----NPRPGSCSVGSPSSDKALTFMKDHFL-----MDEQVVGTFLLVK 411  
QY 428 DRDRMTSVIAYVYK-----NHSIAFVGTGSKLKRVDGPRGNALQVETVQVWDPPGVL 482  
Db 412 SGVEYTRAVETAQGLDGHSHLVWYLGTTTGLSHKAVVSGDSSAHLVEEIQLFDPPEV- 470  
QY 483 RDMAFSKDHEQLYIMGERQLTRVPVESCQYQSCGECGLSGDHPHCGW-----CVLH-- 533  
Db 471 RNQLAPTOGAVFVGSGGVVRPRANCVSVCVDCVLARDPHCAMDPESRTCCLLSAP 530  
QY 534 NCTTRKERCERSEPRFAS-EMKQCVR-----LTVHPNNISVQYNVLLVLETY 582  
Db 531 NLNSWKQDMERGHPWEACASGPMRSRLRPSRPOIIRKVLAVPNSILPELPCPHLSALASY 590  
QY 583 -----NVPESLGVNCTFEDLSMDGLVVGNGIQCVSPAKEVPRIITENG-DHHVVQ 634  
Db 591 YNSHGPAAPVEASSTVYNGSLLILVQDG--VGLYQCNW-----TENGFSYPVIS 638  
635 LQLKSKE 641  
639 YWVDSQD 645  
RESULT 14  
SM4B\_HUMAN STANDARD; PRT; 673 AA.  
ID SM4B\_HUMAN







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 18:51:41 ; Search time 124 Seconds  
(without alignments)  
3150.527 Million cell updates/sec

Title: US-09-964-956-13

Perfect score: 9990

Sequence: 1 MKAMPWNWTCLLSHLLVMG.....QKLYKLEQVITMSLDSNK 1896

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organalle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Description
1	6782	67.9	075051
2	6764	67.7	P70207
3	6393.5	64.0	Q91823 xenopus lae
4	6337	63.4	P70206 mus musculus
5	5984.5	59.9	P70208
6	5847.5	58.5	Q91W2
7	4986	49.9	Q91W1
8	3641	36.4	Q9V491
9	3631	36.3	Q9V491
10	3386.5	33.9	Q96GN9
11	2932.5	29.4	Q96GN9
12	2858	28.6	Q96GN9
13	2841	28.4	Q96GN9
14	2792.5	28.0	Q96GN9
15	2710	27.1	Q96GN9
16	2708	27.1	Q96GN9

17	2707	27.1	2135	4	Q9UIV7
18	2612.5	26.2	1892	11	Q9QY40
19	2583	25.9	1891	4	Q9NNY1
20	2583	25.9	1909	4	Q9HDA4
21	2583	25.9	1912	4	Q9ULI4
22	2580	25.8	1841	4	O15031
23	2475.5	24.8	1985	4	Q9Y4D7
24	2005	20.1	454	4	Q9BRU1
25	1609.5	16.1	1806	5	O45657
26	1449	14.5	328	11	Q8R1I4
27	1374.5	13.8	1574	11	Q9QZC2
28	1373	13.7	258	6	Q9GLW2
29	1349	13.5	1568	4	O60486
30	1324	13.3	540	4	Q9BTQ2
31	1322	13.2	317	4	Q9NTD4
32	1152.5	11.5	470	11	Q991F0
33	1147.5	11.5	470	4	Q9BSU7
34	962	9.6	278	5	Q8T6S7
35	962	9.6	278	5	Q8STG6
36	920	9.2	816	5	Q9V4A7
37	766	7.7	177	4	Q9NSM6
38	678	6.8	729	4	Q9UJ93
39	653	6.5	184	4	Q9HAE7
40	652	6.5	1375	13	Q9W650
41	611	6.1	1382	13	Q90975
42	559	5.6	1375	13	Q9YGM5
43	556	5.6	1425	13	Q9YGM7
44	538.5	5.4	1404	13	Q08757
45	522	5.2	349	4	Q16293

ALIGNMENTS

RESULT 1

ID	075051	PRELIMINARY;	PRT; 1963 AA.
AC	075051;		
DT	01-NOV-1998 (T-EMBLrel. 08, Created)		
DT	01-NOV-1998 (T-EMBLrel. 08, Last sequence update)		
DT	01-JUN-2002 (T-EMBLrel. 21, Last annotation update)		
DE	KIAA0463 protein (Fragment).		
GN	KIAA0463.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RX	MEDLINE=98116662; PubMed=9455484;		
RA	Seki N., Ohira M., Nagase T., Ishikawa K., Miyajima N., Nakajima D.,		
RA	Nomura N., Ohara O.;		
RT	"Characterization of cDNA clones in size-fractionated cDNA libraries		
RT	from human brain."		
RL	DNA Res. 4:345-349(1997).		
DR	EMBL: AB007932; BAA32308.1;		
DR	InterPro; IPR002909; IPT_TIG.		
DR	InterPro; IPR003659; Flexin-like.		
DR	InterPro; IPR002165; Flexin_repeat.		
DR	Pfam; PF01437; PSI; 3.		
DR	Pfam; PF01403; Sema; 1.		
DR	Pfam; PF01833; TIG; 4.		
DR	SMART; SM00429; IPT; 4.		
DR	SMART; SM00423; PSI; 3.		
FT	NON TER		
SQ	SEQUENCE 1963 AA; 218302 MW; 1COAB9B4A35F7804 CRC64;		

Query Match 67.9%; Score 6782; DB 4; Length 1963;  
Best Local Similarity 66.9%; Pred. No. 0;  
Matches 1278; Conservative 238; Mismatches 350; Indels 44; Gaps 9;

QY 23 STLLTRQAPLQKOR-----SFVTRGPAE- 49  
Db 60 SLLT--AAPLSMEORPWRPALEVDSDRSVLLSVVWVLLAPPAAGMPQFSTHSENRD 117  
QY 50 GFNHLVVDERTGHIYLGAVNRIRYKLSADLVVTHETGPDNDPKCPVPPRIVOTCNEPLT 109  
Db 118 TFNHLTVHOCTGAVYVYGAIRNVKLTGNLTQIAHAKTGPEDNKSVPPLIVOPCSEVLT 177  
QY 110 TTNVNMKLLIDYKENRLIACGSLYOGICLKRLLEDLFLGEPVHKKEHYLSGVNESGSV 169  
Db 178 LTNVNMKLLIIDYSENRLLACGSLYQVCKLRLDUDLFLVESHKKEHYLSVNMKTGT 237  
QY 170 FGVISVSNLDDLFATATADGPEYPTTISRKLTKNSBADGNFAVVFHDEFVAMIKI 229  
Db 238 YGIVRSEGEDGLFICTADVGQDYEPTLSSRKLPEDPSSAMLDYELHSDVFSSLIKI 297  
QY 230 PSTFTIIPFDIYVYVGFSSGNFVFLTLOPEMVSPPG---STTKQVYTSKLVRLCKE 286  
Db 298 PSDTALVSHFDIFYIYGFASGGFVFLTVQPE--TPEGVAINSAGDLFTSIRIVRLCKD 355  
QY 287 DTAFNSVVEPIGERSGVYRLLQAAYLKAGAVLGRITLVHPDDLLFTVSKGOKRK 346  
Db 356 DKPFXSVSLPFCGTRAGVYRLLQAAYLAKPGDSLAAQAFNITSQDDVLFAIFSKGQY 415  
QY 347 MKSLDESALCIFILKQINDRIKRLQSCYRGEGTDLAWLKVDIPCSSALLTIDNFCG 406  
Db 416 HHPDDLSALCAPIRAINLQIKRLQSCYQEGNLELWLLGDKVQCTKAPVPIDNFCG 475  
QY 407 LDWNAPLGVSDMVRGIPVFTEDRDMTSVIAVYVKNHSLAFVGTGSKLKIIRVGDGRGN 466  
Db 476 LDINQPLGSGTPVEGLTLYTTRDMTSVSAVYVYGVVFGTSGKLLKIRADGPPHG 535  
QY 467 ALOYETVOVV-DPGVPLRDMAFSDKHOLYIMSERQLTRVPSVSCGOVQSCGCLGSGDP 525  
Db 536 GVOYEMVSVLKGSPILRDMAFSDKHOLYIMSERQVTRVPSVSCGOYTTCGBCCLSGDP 595  
QY 526 HCGWCVLHNTCTRKERCERSEKPRFASEMCKQVRLTVHPNNISVSQYVLLVLETYNVP 585  
Db 596 HCGWCVLHNTCTRKERCERSEKPRFASEMCKQVRLTVHPNNISVSQYVLLVLETYNVP 655  
QY 586 ELGAGVNTCTEDISEMDGLVVGNOIQCYSPAACEVPRIITENGHHVVOIQLSKETGMT 645  
Db 656 DLSAGIACAFGNLTVESGVSGSVICISPGKDPV-VIPLDQDWFGLEQLQSKETGKI 714  
QY 646 FASTSFVYVNCVHNSCLSCVESPYRCHWKYRHCVTDPKTCFQSGRVLKPDCCOLL 705  
Db 715 FVTEFEKYNCSAHLCLSCVNSAFRCHWKYRNLCTHDPKTCFQSGRINISDCCQLV 774  
QY 706 RVDKILVPVEVIRPITLKAKNLPQPSQGRGECILNIQSGEQRVPAALRPNSSVQCOCT 765  
Db 775 PTBEILLIPVEVKPITLKARNLPQPSQGRGECVNLQAGIHRVPAALRPNSSVQCOQS 834  
QY 766 SYSEGMEINNLVPELTVWNGHNINPAQNKVHLKCGAMRESGLCLKADPDPACGW 825  
Db 835 SYQYDGMIDISNLAVDAVWNGNFIIDNPQBLKHLKYCAQRESGLCLKADPDKFCGW 894  
QY 826 CQPGQCTLQCHCPAORSOMLELSGAKSKCTNPRIITEIIPVTGREGTKVIRGENLGL 885  
Db 895 CSGERCTLQCHCTSPSPDLWDSHNVKNSNPQITTEILTVSGPPEGTRVTHGVNLGL 954  
QY 886 EFRDIASHVAVGECPLVDGYTPABQIVCEMGEAKPSQHAGFVEICVAVCRPEFMARS 945  
Db 955 DFSEIAHHVAVGAVPCPTPLPGEYIIAEOIVCEMHALVGTGSGPVLICIGCKEPEFTKS 1014  
QY 946 SOLYFMTLISDLKPSRGPMSGGTQITITGNLACSNVWVMEFKQPCLPFRHSPSYIV 1005  
Db 1015 HQQYTFNVPVSLNIRGPESGGTMTITGHYLGAGSSAVYIGNOTCFEYHSGMSEIV 1074  
QY 1006 C-NTTSDLEVEMKVSQVDRAKHQDLFOYVDPDPTVRIEPEWSIVSGNTPVAVWGT 1064  
Db 1075 CVSPSSNGLGPVPSVSVDRAHVDSNLQFEYIDDPVRIEPEWSIASGHTPLTITGN 1134  
QY 1065 LDLIQNQIRAKGGKSHIINICEVLNATEMTQAPALALGDPDHQSDLTERPEEFGFILDN 1124

Db 1135 LDVIOEPIRVKFKNGESVNVKVVNTTTLTCLAPSLTDTYRPGLDTVRDEDFGVFN 1194  
QY 1125 VQSLIILNKNTFTYVNPVFEAFPGSGIILELKPCTPIILKGNLIPPVAGGNVKNLYTVL 1184  
Db 1195 VQSLIITNDYKFIYFNPFTFELSGPTGVDLQKPSPIILKGNLCPASGG-AKUNYTVL 1253  
QY 1185 VGRKPCVTVTVDVOLLCEPNLGRHKVMARVGMYESPGMVYIAPDSPLSPAIVSIAV 1244  
Db 1254 IGETPCAVTVSETOLLCEPNLTQHKVMHVGMVSPGSGSVISDLSLLTLPVIAVIAA 1313  
QY 1245 AGGLIITIVAVLIAYKRGRESDLTKRLQOMDNLESRLVALECKEPAELQTDIHELT 1304  
Db 1314 GGSLLIIVILVIAKRGRENDLTKRLQOMDNLESRLVALECKEPAELQTDINELT 1373  
QY 1305 SLDGAGIPFLDVTYVTLVLPFGIEDHPVLRDLVEPCYRQERVEKGLKFAQLNNKVP 1364  
Db 1374 SLDGRGIPLDYKTYAMRVLPFGIEDHPVLRDLVEPCYRQERVEKGLKFAQLNNKVP 1433  
QY 1365 LLSFIRLESQSFMRDRGNVASLIMTVLOSLEYATDVLKQLLADLIDKNLESKNHPK 1424  
Db 1434 LLFIRILEQSFMRDRGNVASLIMTVLOSLEYATDVLKQLLADLIDKNLESKNHPK 1493  
QY 1425 LLRRTESVAEKMLTNFTFLYKPLKECAGEPLFSLPCAIIKQOMKEGPIDAITGEARYS 1484  
Db 1494 LLRRTESVAEKMLTNFTFLYKPLKECAGEPLFSLPCAIIKQOMKEGPIDAITGEARYS 1553  
QY 1485 LSEDKLIRQOIKYKTLVLSVSPDNANSPEVPVKILNCDDITOVKEKILDAIFKNVPCSH 1544  
Db 1554 LSEDKLIRQOIKYKTLVLSVSPDNANSPEVPVKILNCDDITOVKEKILDAIFKNVPCSH 1613  
QY 1545 RPKAADMDLEWRQSGARMILQDEDTIKIENDWKRLNTLAHYQVPGDGSVVALSKQVTA 1604  
Db 1614 RPKAADMDLEWRQSGARMILQDEDTIKIENDWKRLNTLAHYQVPGDGSVVALSKQVTA 1673  
QY 1605 YNAVNTSVRTSASKENMIRYTGSPDLSRTPMITPDLESQVGMHVLKQNHGDDOK 1664  
Db 1674 YNAVNTSVRTSASKENMIRYTGSPDLSRTPMITPDLESQVGMHVLKQNHGDDOK 1733  
QY 1665 EGDGRGKMSVSEIYLLRLLATKGTQKQFVDDLPETIFSTAHRSALPLAIKYMFDLDEQA 1724  
Db 1734 EGDGRGKMSVSEIYLLRLLATKGTQKQFVDDLPETIFSTAHRSALPLAIKYMFDLDEQA 1793  
QY 1725 DKGHIDPHVHTWKSNCPLRFVNMINKPQFVFDIHKNSITDACLVSVAQTFMDSCST 1784  
Db 1794 DRHSIHDTVHTWKSNCPLRFVNMINKPQFVFDIHKNSITDACLVSVAQTFMDSCST 1853  
QY 1785 SEHRLGKDSNKLKYADIPSKYKMWERYSDIGKMPAISDQDMNAYLAESRMHNEF 1844  
Db 1854 SEHRLGKDSNKLKYADIPSKYKMWERYSDIGKMPAISDQDMNAYLAESRMHNEF 1913  
QY 1845 NTMSALSEIFSVYKYSBEILGPDHDDQCKQKQKLYAKLEOVITLMSLDS 1894  
Db 1914 NTMSALSEIFSVYKYSBEILGPDHDDQCKQKQKLYAKLEOVITLMSLDS 1963  
RESULT 2  
P70207 PRELIMINARY; PRT; 1884 AA.  
AC P70207;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
GN PLEXIN 2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRIN.  
RX MEDLINE=96400270; PubMed=8806646;



```
Db 1749 FDHKGSIITDACLVSVAQTFMDSCTSEHRLGKDSKSPNKLLYAKDIPSYKNWVERYADI 1808
Qy 1819 GKMPALSDQDMNAYLABQSRMHNFTMSALSIPIFYGVGYSEIILGPLDHDQCKQK 1878
Db 1809 AKUPALSDQDMNAYLABQSRHATEFNLSALNBIYVSVKSEILGALQEQQARRQR 1868
Qy 1879 LAYKLEQVITIMSLDS 1894
Db 1869 LAYKVEHLINAMSTIES 1884

RESULT 3
Q91823
ID Q91823 PRELIMINARY; PRT; 1905 AA.
AC Q91823;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Flexin precursor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95329274; PubMed=7605632;
RA Ohta K., Mitutani A., Kawakami A., Murakami Y., Kasuya Y., Takagi S.,
RA Tanaka H., Fujisawa H.;
RA "Flexin: a novel neuronal cell surface molecule that mediates cell
RT adhesion via a homophilic binding mechanism in the presence of calcium
RT ions.";
RL Neuron 14:1189-1199(1995).
DR ENBL: D38175; BAA07374.1; -.
DR InterPro: IPR002909; IPT_TIG.
DR InterPro: IPR003659; plexin-like.
DR InterPro: IPR002165; plexin_repeat.
DR InterPro: IPR001627; Sema.
DR pfam: PF01437; PSI; 3.
DR pfam: PF01403; Sema; 1.
DR pfam: PF01833; TIG; 4.
DR SMART: SM00429; IPT; 4.
DR SMART: SM00423; PSI; 3.
KW Signal.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 1905 PLEXIN.
SQ SEQUENCE 1905 AA; 214401 MW; 243C059690B33D64 CRC64;

Query Match 64.0%; Score 6393.5; DB 13; Length 1905;
Best Local Similarity 63.7%; Pred. No. 0;
Matches 1222; Conservative 245; Mismatches 396; Indels 55; Gaps 13;

Qy 6 WNWTCILSHLLMVGSGSTLLTROPAPLSQQRQSFVTFRGPAGFNHLVVDERTGHIYL 65
Db 14 WTLVLVLGSIATGDGS-----PKDFTFTGSD-WSLTHLVVHNKITGEVIV 58

Qy 66 GAVNRIYKLSLSLKVLTHTGPDNDPKCYPPRIVOTCNELPTTNNVNNKMLIDYKEN 125
Db 59 GAINRIYKLSNLTLLRTHVTGPDVEDNEKCYPPSPVQSCPHGLITNNVNNKMLIDYSDN 118

Qy 126 RLACGSLYOGICKLLRLDLFKLGPYHKHEHYLGSVNESGSVFGVTVSYSNLDDKLF 185
Db 119 RLACGSASQICQFLRLDDLFLKLGEPHRRKEHYLSSVNESGTVSGVIVPVGQNKLFV 178

Qy 186 ATAVDGRPEVPTTSSRKLTNSADGFAVVFHDEFVSMIKIPSDFTTIIIPFDIYV 245
Db 179 GTPIDGSEYFPTLSSRKLLNGENAEFMFGVYQDEFPVSSQKLPSPDFTLSPFTDIYV 238

Qy 246 YGSSGNNFVPLTLQ--PEWSPPGSTTKQVYTKLVRCKEDTAFNSYVEVPIGERS 303
Db 239 YSFSEQFVYLTILQDQLTQSP--DSTGEQFTSKIVRLCVDDPKFSYVEFPIGCKD 296

Qy 304 GVEYRLLOQAYLSKAGAVLGTGLGVHPPDDDLLFTVFSKGQRKMKSLDESALCIFILKQI 363
```

```
Db 297 GVEYRLLOQAYLSKPGRLAKELGISEDILFTVFSQQRNRIKPPRESVLCLFTLKKI 356
Qy 364 NDRKERLQSCYRCEGTLDLAWLVKDIIPCCSALLTIDDNFCGLDMNAPLGVSDMVRGIP 423
Db 357 KDKIKERLQSCYRGGGKLSPLWLNKELGCLNSPLQIDDNFCGQDFNQLPGTIVIEGTP 416
Qy 424 VFTEDRDMTSVIAYVYKNHSLAFVGTGKGLKRIIVDGPGRNA--LQYETVQVVDGP 480
Db 417 LFLDKEDGMSVAAAYDYGHTVVFAGTSGRVKILVDLSASSHLVQOYENVVHEGNA 476
Qy 481 VLDMAFSKDHEQLYIMSERQLTRVPVESCQYQSCGCLSGDPCGWCVLHNTCTRKE 540
Db 477 ILRDLVLSQRYIYAMTEKQVTRVPVESCQYESCDTCLGSRDPHCGCVLHNMCSRXD 536
Qy 541 RCERSKEPRRFASEMKOCVRLTVHPNNISVSOYNVLLVLTVNVPELSAGVNCCTEDLSE 600
Db 537 KCRADELHRTSDQRCVQVTVHPKNISVTVSEVPWLQAWNVPDLNAGVNCSEDFTE 596
Qy 601 MDGLVVGNIQICYSYPAAKEVPRIITENGHDHVVQLKSKETGMTFASTSFVYNCNVHN 660
Db 597 MEGRIIDGKIYCTSPSAKEVIPITRGHGDKRVVLYLKSKEGKGFASVDFVYNCVSHQ 656
Qy 661 SCLSCVESPYRCHWCKYRHVCTHDPKTCFQGERVKLPEDCPOLLRVDKILVPEVVKI 720
Db 657 SCLSCVNGSPFCHWCKYRHVCTHNAADCSQFGRVNMSEDCPQILPSSQIYIPVGVPKPI 716
Qy 721 TLKAKNLPQPOSGORGVEICILNIQSEORVPALRPNSSVOCQNTSYSEGMENNLVPE 780
Db 717 TLTAKNLPQPOSGORGVEICIFHIPSQVTRVTRVTRVTRVTRVTRVTRVTRVTRVTR 776
Qy 781 LTVVNGHFNIDNPAONKWHLYKCGAMRESCLCLKADPDPFACGWCQGGQCTLRQHCFA 840
Db 777 LSVVNGHFVIDNPNQIQAHLYKCSALRESCLCLKSDRRPFCGWCSEKKTLCRNCPT 836
Qy 841 QESOWLELSGAKSKCTNPRITEIIPVTGPRCGTKVTIRGENLGFEDRIASHVKVAGYE 900
Db 837 LENPMHASTANSRCDPKITKLPETPGCGTSLTITGENLGRFEDIRFGVRVGHVM 896
Qy 901 CSLVDGYIPAEQIVCEMGEA-KPSQAHGAFVEICVAVCRPEPMARSSQLYFMTLTSLD 959
Db 897 CVPVESEVISAQIVCEINDAGRITVEHAQVEVCVKDCSQDYRAISPSTFVLPSEFNV 956
Qy 960 KPSRPGSGGTQVTTITGNLNGSVNVVMFGKQPCIFHRRSPSYVTCNTSSDEVEMLKV 1019
Db 957 TPRSGLSGGTWISTEGNYLNAGSDVSAIGRCPMFSWRTAKEIRCKTPOGPSTGKAEI 1016
Qy 1020 SVQVDRAKIHQDLV-FQYVEDPTIVRIEPEWSIVSGNTPIAVMGTHLDLIQNPQIRAKHG 1078
Db 1017 QILINRATMNSVHYNYTETPTQKIEPWSIASGGTPLIVTGMNLATIKPKIRAKYG 1076
Qy 1079 GKEHINICEVLNATEMTQAPAL---ALGPPDHQSDLTBRPERFGLDNVQSLLILNKTN 1135
Db 1077 DVEKENNCTLYNDITWVCLAPSDVNDPLRSPPENGD---RPDEIGFIMDNVHALLIVNTTS 1133
Qy 1136 FTYVPNVFEAFSGSILELPGTPIILKKNVLIIPVAGGNVKNLYTVLVGSKPCVTTVTS 1195
Db 1134 FLYYVDPVFEPLTAGNLKPSPLIIGRNLII-PAAPAGNFRLNVTVLIGTTPCALTVS 1192
Qy 1196 DVOLLCEPNLIGBHKWARGVGMYSVPMVVIAPDSPLSIPAIVSIAVAGGLLIIPVA 1255
Db 1193 ETQLLCEPNLTGQHKVTKAGGPEYSPETLQYSDSLTLTLPALIGGGGGLLIILII 1252
Qy 1256 VLIAYKRKRSRSDTLTKRLQMDNLSRVALECKEFAELQTDIHELTDLDGAGIPFL 1315
Db 1253 VLIAYKRKSRDRTLTKRLQMDNLSRVALECKEFAELQTDIHELTDLDGAGIPFL 1312
Qy 1316 DRYTYTRVLPFGIEDHVLRLDLVPGYQQRVEKGLKLPALQILNNKVFLLFIRPLESQ 1375
Db 1313 EYRTYTRVLPFGIEDHVPVKMEV---QANVEKSLTLFGQLLTKKHFLLFIRLEAQ 1368
Qy 1376 RSFMRDRGNVASLMTVLOSLEKLVATDVLKOLLADLIDKNLESKNHPLKLLRRRTESVAE 1435
```

Db 1369 RFSMRDRGNVSLINTALQGEYATGVLKQLLSLIEKNLESKNHPKLLLRTERESVAE 1428  
QY 1436 KMLTNWFTLLYKFLKECAGEPLSFLCAIKQOMKEGPDIDAITGEARYSLSEDKLIRQOI 1495  
Db 1429 KMLTNWFTLLYKFLKECAGEPLFMLHCAIKQOMKEGPDIDAITGEARYSLSEDKLIRQOI 1488  
QY 1496 DYKTL-----VLSCVSPONANSPVVPVKILNCDDTITQVKEKILDAI 1536  
Db 1489 DYKTLNFCADDVGLSDSCCRSPQTLNVCNPNENAPEIPVKVLCNDTITQVKEKILDAV 1548  
QY 1537 FKNVPCSHPKADMDLEWRQSGAEMILQDEDTTKIENDWKRIANTLAHYQVPCGSVA 1596  
Db 1549 YXGVPSYQPKAGMDLEWRQSGAKIILQDEDDVTTKINDWKRLNLAHYQVTDGSSVA 1608  
QY 1597 LVSKQVTAYNANVNSTVSRSTASAKYENMIRYTGSPDSLSRSTPMITPDLESYGKVMHLVK 1656  
Db 1609 LVPKQNSAYNISNSSTFTK-SLSRYESMLRTASSPDSLSRSTPMITPDLESYGKVMHLVK 1667  
QY 1657 NHEHGDQEGDRGSKMSEIYLTRLATKGTLOKFVDDIFETIFSTAHRSALPLAIKYM 1716  
Db 1668 NHDHLDQREGDRGSKMSEIYLTRLATKGTLOKFVDDIFETIFSTAHRSALPLAIKYM 1727  
QY 1717 FDFLEQADKHGHDPVHTWKSNCPLRFWVNMINKPOFVEDIHKNSITDACLSSVAQ 1776  
Db 1728 FDFLEQADKHGHDPVHTWKSNCPLRFWVNMINKPOFVEDIHKNSITDACLSSVAQ 1787  
QY 1777 TFWDSCTSEHLRGKDPSPNKLLYAKDIPSYKNWERYYSIDIKMPAISDQDMNAYLAEQ 1836  
Db 1788 TFWDSCTSEHLRGKDPSPNKLLYAKDIPSYKNWERYYSIDIKMPAISDQDMNAYLAEQ 1847  
QY 1837 SRMHNEFTMSALSIFSYGVKYSIEILGPLDHPDQCGKQKLAYKLEQVITLMSLDS 1894  
Db 1848 SRLHLQSFMSALHEIYSVITKYRDEILTALEKDSQARRQLRSKLEQVIDTMAQSS 1905

RESULT 4  
P70206 PRELIMINARY; PRT; 1894 AA.  
ID P70206  
AC P70206;  
DT 01-FEB-1997 (TREMBlrel. 02, Created)  
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Plexin 1.  
GN PLXN1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=96400291; PubMed=8806667;  
RA Kameyama T., Murakami Y., Suto F., Kawakami A., Takagi S., Hirata T.,  
RA Fujisawa H.;  
RT Identification of a neuronal cell surface molecule, plexin, in  
RT mice.";  
RL Biochem. Biophys. Res. Commun. 226:524-529(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=20191894; PubMed=10725340;  
RA Shimizu M., Murakami Y., Suto F., Fujisawa H.;  
RA "Determination of Cell Adhesion Sites of Neuropilin-1.";  
RL J. Cell Biol. 148:1283-1294(2000).  
DR EMBL; D86948; BAA13188.1; -.  
DR MGD; MGI:107685; Plxn1.  
DR InterPro; IPR002909; IPT TIG.  
DR InterPro; IPR003659; Plexin-like.  
DR InterPro; IPR002165; Plexin\_repeat.  
DR InterPro; IPR001627; Sema.  
DR Pfam; PF01437; PSI; 3.  
DR Pfam; PF01403; Sema; 1.  
DR Pfam; PF01833; TIG; 4.

DR SMART; SM00429; IPT; 4.  
DR SMART; SM00423; PSI; 3.  
SQ SEQUENCE 1894 AA; 211098 MW; A856BB29C6824C94 CRC64;  
Query Match 63.4%; Score 6337; DB 11; Length 1894;  
Best Local Similarity 63.5%; Pred. No. 0;  
Matches 1209; Conservative 267; Mismatches 408; Indels 20; Gaps 13;  
QY 1 MKAMPNWTCLLSHLLMVGMSSTLLTRQAPLTSQKORSFVTFRGEPAEGFNLHVVDERT 60  
Db 1 MPPLPSSRTLLLLLLLLRGVWTAISSPPAGLG-PQPAFTTFVASD-WGLTHLVVHQOT 58  
QY 61 GHYLGAVNRIYKLSDDLKVLVTHETGPDENPKYCPPIVQTCNEPIITNNVNMKLLI 120  
Db 59 GEVTVGAVNRIYKLSGNLTLLRAHVTCGVEDNEKYPSPSVQSCPHGLGSDNVNKLILL 118  
QY 121 DYKXNRIIAGSLYQGIKLLRLLEDLPKLGEPYKKEHYLSGVNESGVSFGVIVS--YSN 178  
Db 119 DYANRLIACGSASQGIQFLRLDDLPKLGEPHKKHYLSVREAGSMAGVLIAAGPQG 178  
QY 179 LDDKLFATADVQKPEYPTTISRKLTKNSADGMEAVVFHDEFVASKIKTPSDFTTIP 238  
Db 179 GOAKLFVGTIDGKSEYPTTISRRLMANEDAMFGVYQDEFVSSQKLPISDTLSKFP 238  
QY 239 DFDIYVYVYFSSGNFVYFLTLQ--PEMVSPPGSTTKQVYTSKLVRLCKEDTAFNSYVEV 296  
Db 239 AFDIYVYVYFRSEQFVYVLTQLDQLTSP--DAAGEHFFTSKIVRLCNDPKFYVYVEF 296  
QY 297 PIGERSGVYRLLOAYLSKAGAVLGRTLGVHPPDDLLFTVFSKQGRKMKSLDESALC 356  
Db 297 PIGCEQAGVYRLVQDAYLSRPGQALAKQLAEDEEVLTFTVFAQGGQKRVKPKKESALC 356  
QY 357 IFILKQINDIRKELQSCYRGEGLDLAWLKVKDI PCSSALLTTDDNFCEGLDMNAPLGV 416  
Db 357 LFTLRAIKERIKQSCYRGEGLSLPMLLNKLCGINSPLQDDDDFCQDQFNQPLGT 416  
QY 417 DMVRGIPVFTEDRMTSVIAYVYKNSHLAFVGTGSKGLKIRVD--GPRGN-ALQYETV 473  
Db 417 VTIEGTFPLVFDKEDGLTAAAYDYQGRTVVTFAGTRSGRIKILVDLANPSGRPALAYESV 476  
QY 474 QVVDPGVLRDMAPSKDHEQLYIMSERQLTRVPVSCGQYOSCGECIGSGDPHCGWCVLH 533  
Db 477 VAQEGNPILRDLSPNRQYLYAMTERQVTPVPEVCYQVTSCELCGLSGRDPHCGWCVLH 536  
QY 534 NTCTRKERCERSKEPRFASMEKOCVRLTVHPNNISVSQYVNVLLVETVYVPELSAGVNC 593  
Db 537 SICSRQACRABEPQRFASDLQCVLTQVPRNVSVTMSQVPLVQANVPDLSAGVNC 596  
QY 594 TFDLSEMDGLVVGNOIQCYSPAKEYPRIITENGDDHVVQLQKSKETGMTFASTSFVP 653  
Db 597 SPEDFTETESILEDGRVHCHSPSAREVAPITQGGQDQVVKLYLKSKEKGFASVDVFP 656  
QY 654 YNCSVHNSCLSVESPYRCHWKYRVHCTHDPTCTCSFOEGRVVKLPEDCPQLLRVDKILVP 713  
Db 657 YNCSVHOSCLACVNGSPFCHWKYRVHCTNNAADCAFLGRVNMSEDCPQLLPSTHIYVP 716  
QY 714 VEVIKPTTLAKNLPQPSGORGVECTLNTGSEORVPALRFNSSVOCQNTSVSYEGWE 773  
Db 717 VGVPKPTTLAARNLPQPSGORGVECTLNTGSEORVPALRFNSSVOCQNTSVSYEGND 776  
QY 774 INNLVELTVVWNGHFNIDNPAQNKVHLKCGANRESGLCLKADPPDFACGWCQPGQCT 833  
Db 777 VSDLPVNLVWNGNFVIDNPQIAHLKCPALRQSCGLCLKADPPRECWCVAERCS 836  
QY 834 LROHCPAQE--SQMLESGLAKSKCTNPRITEIIPVTPREGGTVKTVIRENIGLEPRDIAS 892  
Db 837 LRHCPADSPASWHAHSGSRCTDPKILKLSPTGPRQGGTRITITGENIGLRFEDVRL 896  
QY 893 HVKVAGVCECPLVDGYTPAQIVCEMGEAKPSQ--HAGFVEICVAVCRPEEFMARSSQLYVF 951  
Db 897 GVHVGVKLCSPVSEYISAEQIVCEIDASTLRAHDALVEVCVRDCSLHYRALSPRRFTF 956  
QY 952 MTLTSLDKPSRGPMSGGTQVTTGTNLNAGSNVVMFGKQPCFLFHRSPSYIVCNTTSS 1011

Db 957 VTPFVYVSPSRGSLGGTWIGGSHNAGSDVAVSIGRCPSCFSWRNSREIRCLTPPG 1016  
QY 1012 DEVLEMKSVQVDRAKI-HODLVPOYVEDPTVIRIEPWSIVSGNTPIAVWGTHLDLION 1070  
Db 1017 HTPGSAPVIVINRAQLNSPVEKYNTEPTILRIDPEWSINGGTLTTLVTGINLATVRE 1076  
QY 1071 PQIRAGHGKHEHINI CEVLNATEMTCAPALALGPDHQSOLDTERPEFGFILDNVQSLLI 1130  
Db 1077 PRIRAKYGGIERNSCMVYNDITWCEAPSIDNPKRSPPELGERPDIGIFIMONVRLTV 1136  
QY 1131 LNKNTFYYPNPVEAFPGSGIILELKPPTIILKGNLIPPVAGNVKLYNTVLVGBKPC 1190  
Db 1137 LNSSFLYYPDPVLEPLSPGLLEKPSLILKGNLILPP-APGNSRLNYTVLIGSTPC 1195  
QY 1191 TTVTSVDOLLCESPNLIGRHKWARVCGMEYSYSGMVYIAPDSPLSLPAIYSIAVAGLLI 1250  
Db 1196 ILTVSETQLLCEAPNLTGQHKVTVRAGGFBSFGMLQVYDSLLTLPFAI VIGGGGGLLL 1255  
QY 1251 IFIVAVLIAVYKRSRSDTLKRLQMDNLESVALLEKEAFELQTDIHELTSDLGA 1310  
Db 1256 LVIVAVLIAVYKRSRSDRTLKRLQMDNLESVALLEKEAFELQTDIHELTSDLGA 1315  
QY 1311 GIPFLDYRTVMRVLPGIEDHPVIRLEVPYQERVEKGLKLPALINNKVFLLSFIR 1370  
Db 1316 GIPFLDYRTVMRVLPGIEDHPVIRLEVPYQERVEKGLKLPALINNKVFLLSFIR 1371  
QY 1371 TLESQSFMSRDRGNVSLTMTVLOSLEKLVATDVLKOLLADILDKNESKNHPKLLLRRT 1430  
Db 1372 TLESQSFMSRDRGNVSLTMTVLOSLEKLVATDVLKOLLADILDKNESKNHPKLLLRRT 1431  
QY 1431 ESVAEKMNTWFTLLYKFLKECAGEPLSFCAIKQOMEKGPIDAITGEARYSLSEDKL 1490  
Db 1432 ESVAEKMNTWFTLLYKFLKECAGEPLFMYLCAIKQOMEKGPIDAITGEARYSLSEDKL 1491  
QY 1491 IRQOIDDYKTLVLCVSPDNANSPPVVKILNCDITITQVKEKILDAIFKNVPCSHRPAAD 1550  
Db 1492 IRQOIDDYKTLVLCVSPDNANSPPVVKILNCDITITQVKEKILDAIFKNVPCSHRPAAD 1551  
QY 1551 MDLEWROGSGARMTLODEDITTKIENDWKRLNTLAHYQVDPGSVVVALVSKQVTAYNVN 1610  
Db 1552 MDLEWROGSGARMTLODEDITTKIENDWKRLNTLAHYQVDPGSVVVALVSKQVTAYNVN 1611  
QY 1611 STVSRTSASKYENMIRYTGSPDSLSRTPMITPDLESQVXWHVLVKNHEHGDQKGBRGS 1670  
Db 1612 STFTK-SLSRYESMLRTASSPDSLSRTPMITPDLESQVXWHVLVKNHEHGDQKGBRGS 1670  
QY 1671 KMWSEIYLTRLLATKGTLOKVDLLETIFSTAHRSALPLAIKYMDFLDEQADKHQIH 1730  
Db 1671 KMWSEIYLTRLLATKGTLOKVDLLETIFSTAHRSALPLAIKYMDFLDEQADKHQIH 1730  
QY 1731 DPHVHTWKSNCPLRFVWVWIKNPOQVFDIHKNSITDACL SVVAQT FWDSCSTSEHRLG 1790  
Db 1731 DSDVHTWKSNCPLRFVWVWIKNPOQVFDIHKNSITDACL SVVAQT FWDSCSTSEHRLG 1790  
QY 1791 KDSPSNKLLYAKDIPSYKNWERYYSYDCKMPAISDQMDNAYLAPQSRHNMFEPTMSAL 1850  
Db 1791 KDSPSNKLLYAKDIPSYKNWERYYSYDCKMPAISDQMDNAYLAPQSRHNMFEPTMSAL 1850  
QY 1851 SEIPSYGVKYSIEILGPDHDDQCKQKLYLKEQVITLMSLDS 1894  
Db 1851 HEIYSYIAKYKDEILVALEKDEQARRQRLSKLEQVDMTALSS 1894

## RESULT 5

P70208 PRELIMINARY; PRT; 1872 AA.  
AC P70208;  
DT 01-FEB-1997 (Tremblrel. 02, Created)  
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Plexin 3.  
GN PLXN3.

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=96400270; PubMed=8806646;  
RA Kameyama T., Murakami Y., Suto F., Kawakami A., Takagi S., Hirata T.,  
RA Fujisawa H.;  
RT "Identification of plexin family molecules in Mice."  
RL Biochem. Biophys. Res. Commun. 226:396-402(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=10191894; PubMed=10725340;  
RA Shimizu M., Murakami Y., Suto F., Fujisawa H.;  
RT "Determination of Cell Adhesion Sites of Neuropilin-1";  
RL J. Cell Biol. 148:1283-1294(2000).  
DR EMBL; D86950; BAA13190.1; -;  
DR MGI; 107683; Plxn3.  
DR InterPro; IPR002909; IPT\_TIG.  
DR InterPro; IPR003659; Plexin-like.  
DR InterPro; IPR002165; Plexin\_repeat.  
DR InterPro; IPR001627; Sema.  
DR Pfam; PF01437; PSI; 3.  
DR Pfam; PF01403; Sema; 1.  
DR Pfam; PF01833; TIG; 4.  
DR SMART; SM00429; IPT; 4.  
DR SMART; SM00423; PSI; 3.  
SQ SEQUENCE 1872 AA; 207941 MW; 39D5789DCC4830A CRC64;

Query Match 59.9%; Score 5984.5; DB 11; Length 1872;

Best Local Similarity 59.7%; Pred. No. 0;

Matches 1130; Conservative 290; Mismatches 441; Indels 31; Gaps 13;

QY 10 CLLSHLLM-VG-MGSSLTLLTRQAPLQKQSFVTERGEPAEGFNHLVVDERTHYILG 66  
Db 5 CLLPLLFTTIGCGLGSS-----PPRTFVVTDTT-LTHLAVHRTGEVFG 49  
QY 67 AVNRYKLSLDLKVLTHTHETGDEPNKCYPPRIVQTCNEPLTTNNVKNMLLIDYKENR 126  
Db 50 AVNRVFKLAPNLTELRAHVTGPIEDNARCYPPPSMRVCSHRLVPVDNVNKLIDYAAAR 109  
QY 127 LIACSLYQGIKRLRLDLFKLGEPIYKHEHLSGVNMSGVGVIVSVNLDLKLFA 186  
Db 110 LVACGSIWQIGICQFLRLDLFKLGEPIYKHEHLSGVNMSGVGVIVSVNLDLKLFA 169  
QY 187 TAVDCKPEVFTTISRKLTKNSEADGMFAYVPHDEFVASMIKIPSDTFTIIPDFIYVY 246  
Db 170 TAVDCKSEVFTLSSRKLIDEDSGDMFSLVQDEFVSSQIKIPSDTSLYPAFDIYVY 229  
QY 247 GFSSGNFYFTLOPEWVSPPGSTTKQVYTKLVLKEDTAFNSVVEVPICGERSGVE 306  
Db 230 GFVSASFVFLTLQDLTQTLDTAGKEFTSKIVRMACAGDSEFYSEVPEPIGCSWRGVE 289  
QY 307 YLLQAAVLSKAGVIGRLTVGHVDDLLFTVFSKQKQKMKSLDESALCIFILKQINDR 366  
Db 290 YRLVQSAHLAKPELLAQAIGVPADEDLVFTFISQCKNRPANPPROTILCLFLSSINAH 349  
QY 367 IKERLQSCYRGEGTLDLAWLKVDIPCSSALLTIDNFCGLDMNAPLGVSDMVRGIFVFT 426  
Db 350 IRRRIQSCYRGEGTLPALPWLNLKELPCINTPLQINGFCGLVNLQPLGLHVEGLPLLA 409  
QY 427 EDRDRMTSVIAYVYKNSLAFVGTGSKLKKIRVDGPRGNALQYETVQVVDPGVLDMA 486  
Db 410 DSTDGMASVAAVYTHQHSVVFITGRGNLKKRVDSQ-DAQLYETVSVVQSGFILDL 468  
QY 487 FSKDEHQLYIMSERQLTRVPVSGQYQSGCEGLSGDPCGCVLHNTCTRKRCERSK 546  
Db 469 FSPDRHRYLLSEKQVSLPVECTEQLYSCAACIGSGDPCGCVLHNTCTRKRCERSK 528  
QY 547 EPRRFASEMKQCVRLTVHPNNISVSQYNVLLVLETYNVPELSAGVNTCTFDLSEMDGLV 606

Db 529 APHGFAELSKCIQVRNENSVTSQVLTAMENVPDLSVGVSCSFEVTESEAILL 588  
QY 607 -GNOIQCYGPAKAEVPRITITENGDEHVVOLQLKSKETGTMFTASTSFVYNCVSHNSCLSC 665  
Db 589 PSGLRCPSFSLQELQTLTRGHGATHVRLQLLSMTGVRFGADVFFVYNCALQSCMSC 648  
QY 666 VSPYRCHWKYRHHVTHDPKTCFQGRVYKLPEDCPQLLRVDKILVPVEVVKPITLAK 725  
Db 649 VGSPPYCHWKYRHHVTHDPKTCFQGRVYKLPEDCPQLLRVDKILVPVEVVKPITLAK 708  
QY 726 NLPQSGQGRGYECILNIQSGEORVPALRPNSSVOCQNTSYSGEWEINNLPEVITVW 785  
Db 709 NLPQSGQGRGYECVVRVQGRHVRPAVRFNSSVOCQNASYFEGDEFGDTLDFSVW 768  
QY 786 NGHFNIDNPAQNKVHLKYCGMRESGCLCKADDPACQWCGQPGCTLRHCHPAQESOW 845  
Db 769 DGFDPIDKPPFRALLYKYKWAQPSGCLCKADPRNCWCISEHRCQLRAHCPAPKSNW 828  
QY 846 LELSGAKSKCTNPRITEIIPVTPREGGKVTIRGENGLGLEFRDIASHVKVAGVECSPLV 905  
Db 829 MHPQKQKARSHPRITQIHPLTGPKEGGTRVITVGENGLGTSREVG--LRVAGVRCNSIP 886  
QY 906 DGYIPAEQIVCEMGEA-KESQHAGFVEICVAVCRPEFMASSOLYYPMLTILSDLRPSRG 964  
Db 887 TEYVSAERIVCEMEESLSPSPGPAELCVGDCSADFRTQSOQLYFVTPPTDRVSPSRG 946  
QY 965 PMSGGTGVITGTNLNAGSNVVMFCQKPLFRRSPSYVVC--NTTSSDVLKMKYSVQV 1023  
Db 947 PASGGTGLTIGSLDAGSRVVIIRDEGQFVRRDAEIVCISPVSTLGPSPITLAI 1006  
QY 1024 DRAKI-HODLVQFYVEDPTVIRIEPWSIVSGNTPIAVMGTHLDLQNPQIRAKHGKSH 1082  
Db 1007 DHANISNTGVITYTQDPTVTHLEPTWSIINGSTSVTSSTHLLTQEPVRVAKYRGIE 1066  
QY 1083 INICEVLAETMTCPALALGPDHOSDTERPEBPGFILDNVQSLILINKNTFTYYPNP 1142  
Db 1067 TWTQVINDTAMLCAPGIFLGHQPRAGQEHDFEGLFDHVAARSLSRSFTYYPDP 1126  
QY 1143 VFEAFGSGILELKPCTPIILKGNLIPVAGNVKLVTLVNGEKPCVTVSDVOLLCE 1202  
Db 1127 SEFPGPSGLVDKPGSHVVLKGNLIPAAAGSS-RLNTVTLVGGQPCALTUVDTLQCD 1185  
QY 1203 SPNLIGRHKVMARVGMESYSPGVYIAPDPLPAIVSIAVAGGLLIPIFVAVLIAYKR 1262  
Db 1186 SFSQTRQPMVVLVGLFELWGLTHITADRALTPAMVGLAAGGLLAILTIVLVAYKR 1245  
QY 1263 KRESDLTKRLQOMDNLESRALECKEAFELQTDIHELTSDDLQAGIPPLDYRTYM 1322  
Db 1246 KTQADRLTKRLQOMDNLESRALECKEAFELQTDIHELTHMDGVQIPLDYRTYAV 1305  
QY 1323 RVLFGIEHPVLRDLVPGYRQERVEKGLKPAQLINNKVFLLSFIRTLQESRFSMRD 1382  
Db 1306 RVLFGIEHPVLRDLVPGYRQERVEKGLKPAQLINNKVFLLSFIRTLQESRFSMRD 1361  
QY 1383 RGNVASLMTVLQSKLEYATDVLKOLLADLIDKNLESKNHPKLLLRRTESVAEKMLTNWF 1442  
Db 1362 RGTVASLTWVALQSRLDYATGLLKOLLADLIEKNLESKNHPKLLLRRTESVAEKMLTNWF 1421  
QY 1443 TFLLYKFLKECAGEPLFSFCAIKQOMKEGPIDAITGEARYSLSEDKLIRQQIDYKTLVL 1502  
Db 1422 TFLHKLFLKECAGEPLFLYCAIKQOMKEGPIDAITGEARYSLSEDKLIRQQIDYKTLVL 1481  
QY 1503 SCVSPDNANSPEVPVKILNCDITTOVEKITLDAIFKNVPCSHPEKADMDLEWROSGAR 1562  
Db 1482 HVCPESEGAQPVVKVLCNDSITQAKDLDTVYIGIPVSPQRPKAEKMDLEWROSGAR 1541  
QY 1563 MILQDEIDITTKIENDWKRLNTLAHYQVPGSVVALVSQVTAYNVNNSTVRSATSKYE 1622  
Db 1542 IILQDEIDITTKIENDWKRLNTLAHYQVPGSVVALVSQVTAYNVNNSTVRSATSKYE 1600  
QY 1623 NMIRYTGSPDSLRSPMTPTDLESQVGMVHLVKNHEHGQKEDGRGSKMVSIIYLTRLL 1682

Db 1601 SLRAASSPDLRSRAMPRLTPDQAGTKMLHVLVRNHDHTRHREGDRGSKMVSEIYLTRLL 1660  
QY 1683 ATKGTLOKQFVDDLPETIFSTAHRSALPLAIKYMFDLDEQADKHGHDHVRHTWKNC 1742  
Db 1661 ATKGTLOKQFVDDLPETIFSTAHRSALPLAIKYMFDLDEQADKHGHDHVRHTWKNC 1720  
QY 1743 LPLRFVWVMKINQFVDFDIHKNSTIDACLSVVAQTFMDSCSTSEHRLGKDSPSNKLKYAK 1802  
Db 1721 LPLRFVWVMKINQFVDFDIHKNSTIDACLSVVAQTFMDSCSTSEHRLGKDSPSNKLKYAK 1780  
QY 1803 DIPSYKMWERYYSIDICKMPAISODMNAVLAEOSRMHNEFNTMSALSELSEFSYKYSX 1862  
Db 1781 DIPSYKMWERYYSIDICKMPAISODMNAVLAEOSRMHNEFNTMSALSELSEFSYKYSX 1840  
QY 1863 EILGPLDHDQCGKQKLAYKLEQVITLMSLDS 1894  
Db 1841 EILGPLDHDQCGKQKLAYKLEQVITLMSLDS 1872  
RESULT 6  
Q9UIW2 PRELIMINARY; PRT; 1754 AA.  
AC Q9UIW2 PRELIMINARY; PRT; 1754 AA.  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE NOV/plexin-A1 protein (Fragment).  
GN NOV/PLEXIN-A1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N. A.  
RC TISSUE=SKELTAL MUSCLE;  
RX MEDLINE=96149362; PubMed=8570614;  
RA Maestrini E., Tamagnone L., Longati P., Cremona O., Gulisano M.,  
RA Bione S., Tamani F., Neel B.G., Toniolo D., Comoglio P.M.;  
RA "A family of transmembrane proteins with homology to the MET-hepatocyte  
RT growth factor receptor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:674-678 (1996).  
RN [2]  
RP SEQUENCE FROM N. A.  
RC TISSUE=SKELTAL MUSCLE;  
RX MEDLINE=99449305; PubMed=10520995;  
RA Tamagnone L., Artigiani S., Chen H., He Z., Ming G.L., Song H.L.,  
RA Chedotal A., Winberg M.L., Goodman C.S., Poo M.M., Tessier-Lavigne M.,  
RA Comoglio P.M.;  
RT "Plexins are a large family of receptors for transmembrane, secreted  
RL and GPI-anchored semaphorins in vertebrates.";  
RL Cell 99:71-80 (1999).  
DR EMBL: X87832; CAB57274.1; .  
DR InterPro; IPR002909; IPT\_TIG.  
DR InterPro; IPR003659; Plexin-like.  
DR InterPro; IPR002165; Plexin-repeat.  
DR InterPro; IPR001627; Sema.  
DR Pfam; PF01437; PSI; 3.  
DR Pfam; PF01403; Sema; 1.  
DR Pfam; PF01833; TIG; 4.  
DR SMART; SM00429; IPT; 4.  
DR SMART; SM00423; PSI; 3.  
FT NON\_TER 1 1  
FT NON\_TER 1754 1754  
SQ SEQUENCE 1754 AA; 194805 MW; 4F61F0DE86C3D9AB CRC64;  
Query Match 58.5%; Score 5847.5; DB 4; Length 1754;  
Best Local Similarity 63.7%; Pred. No. 0;  
Matches 1114; Conservative 237; Mismatches 378; Indels 19; Gaps 12;  
QY 37 QRSFVTFGEFPAEFGNHLVVDERTGHIYGLGAVNRIYKLSDDLKVLVTHETGPDENPKCY 96  
Db 16 QPPFRTFSGSD-WGLTHLLVHEQTGEVYGVAVNRIYKLSGNLTLLRAHVTFGVEDNEKCY 74

QY	97	P	P	R	I	V	O	T	C	N	E	P	L	T	T	N	N	N	K	L	I	D	I	Y	K	E	N	R	L	I	A	C	G	S	L	Y	Q	G	I	C	K	L	R	E	D	L	F	K	L	G	S	P	Y	H	K	156		
Db	75	P	P	X	V	O	S	C	P	H	L	G	N	T	D	N	V	K	L	L	L	D	Y	A	N	R	L	L	A	C	G	S	A	S	O	G	I	C	O	S	L	R	L	D	X	L	F	K	L	G	S	P	H	R	K	134		
QY	157	E	H	L	S	G	N	B	S	G	S	V	F	G	I	V	S	--	Y	S	N	L	D	K	L	F	I	A	T	A	V	D	G	K	E	P	E	P	T	T	S	R	K	L	T	N	S	E	A	D	G	M	F	214				
Db	135	E	H	L	S	S	V	O	E	A	G	S	M	A	G	V	L	I	A	G	P	P	O	G	O	A	K	L	F	V	G	T	P	I	D	G	K	E	Y	F	P	T	S	R	R	L	M	A	N	E	E	D	A	M	F	194		
QY	215	A	Y	V	P	H	E	F	V	A	S	M	I	K	I	P	S	D	T	T	I	P	D	P	I	I	V	V	V	G	F	S	G	N	F	V	F	L	T	Q	--	P	E	M	S	P	P	G	S	T	T	K	272					
Db	195	G	F	V	Q	D	E	F	S	S	O	L	I	P	E	D	T	L	K	E	P	A	D	I	I	V	V	S	F	R	S	E	Q	F	V	Y	L	T	Q	L	T	S	P	--	D	A	A	G	252									
QY	273	E	Q	V	T	S	K	L	R	L	C	K	E	D	A	F	N	S	V	E	V	P	I	C	E	R	S	G	E	V	E	R	L	L	O	A	Y	L	S	K	A	G	A	V	L	G	R	T	L	G	V	H	P	D	332			
Db	253	E	H	F	T	S	K	I	V	L	C	V	D	D	P	K	F	Y	S	V	E	F	P	I	C	E	Q	A	G	V	E	R	L	V	Q	D	A	Y	L	S	R	P	G	R	A	L	A	H	O	L	G	L	A	E	312			
QY	333	D	L	I	F	T	S	K	O	K	R	M	K	S	L	D	S	A	L	C	I	F	I	L	K	O	I	N	D	R	I	K	E	R	I	Q	S	C	V	R	G	B	T	I	D	L	A	M	L	K	V	D	I	P	392			
Db	313	D	V	L	F	T	P	A	O	Q	K	N	R	V	K	P	E	S	A	L	C	L	F	T	R	A	I	K	E	I	K	E	R	I	Q	S	C	V	R	G	E	G	K	L	S	P	L	M	L	N	K	E	I	G	372			
QY	393	C	S	S	A	L	L	I	D	N	F	C	G	L	D	M	N	A	P	L	G	V	S	D	M	V	R	G	I	P	V	T	E	D	R	M	T	S	V	I	A	Y	V	K	N	H	S	L	A	F	V	G	T	K	452			
Db	373	C	I	N	S	P	L	O	D	D	D	F	R	Q	D	F	N	Q	P	L	G	T	V	I	E	G	T	P	L	F	V	D	K	D	G	L	T	A	V	A	A	D	Y	R	G	R	T	V	F	A	G	T	S	432				
QY	453	G	K	L	K	I	R	V	D	--	G	P	R	G	N	--	A	L	O	Y	E	T	V	Q	V	V	D	P	F	V	L	R	D	M	A	F	S	K	O	H	E	O	L	Y	I	M	S	E	R	Q	L	T	R	V	P	V	S	509
Db	433	G	R	I	K	L	V	D	L	S	N	P	G	R	P	A	L	A	Y	E	S	V	A	Q	E	S	P	I	L	R	D	L	V	L	S	N	H	O	Y	L	V	A	M	T	E	K	Q	V	T	R	V	P	V					

Db	1153	NLUPP-APGNSRLNYTLVIGSTPCTIVTSETQLCEAPNLTGQHKVTVRAGGFSPGTL	1211
Qy	1227	YIAPDSPLSIPAIVSIAVAGGLIIPFIVAVLIAYKKSRESDLTKRLQOMDNLESRA	1286
Db	1212	QVYSDSILLTLPAIVGIGGGGLLLVIVAVLIAYKKSADARTKRLQLQMDNLESRA	1271
Qy	1287	LECKEFAELQTDIHELTSDLGAGIPFDIDRYTMRVLPFGIEDHPVLRDLVEVPGVQE	1346
Db	1272	LECKEFAELQTDIHELTDNDLQAGIPFDIDRYTMRVLPFGIEDHPVLMKMEV----	QA 1327
Qy	1347	RVKSGKLKLAQLINNKFVLLSFIRTIQESQSMRDGNVASLIMTVLOSLEYATDVLK	1406
Db	1328	NVZKSITLFCQLTKKHFLTFIIRTEAQRFSFMRDGNVASLIMTALQEMEYATGVLK	1387
Qy	1407	QLLAADLIDKNLESKNHPKLLRRTESVAEKMLTNWTFLLYKFKECAGEPLSLCAIK	1466
Db	1388	QLLSDLITEKNLESKNHPKLLRRTESVAEKMLTNWTFLLYKFKECAGEPLFMYCAIK	1447
Qy	1467	QOMEKGPIDAITGEARYSLEDKLIQQIDYKTLVLSVSPDNANSPEVPVKILNCDDTIT	1526
Db	1448	QOMEKGPIDAITGEARYSLEDKLIQXIDYKTLVLCVNPENAPPEVPVKGLDCTVT	1507
Qy	1527	QVKEKILDAIFKNVPCSHRPAADMLEWRQSGARMLODEITTKIENDWKRNLTLAH	1586
Db	1508	QAKEKLLDAAYKGVPSQRPAADMLEWRQSGRMARIILQDEDTVTKIDNDWKRNLTLAH	1567
Qy	1587	YQVPDGSVALVSKQVAYNAVNSTVSRTSASKYENMIRYTGSPDSLSRSTPMITPDLE	1646
Db	1568	YQVTDGSSVALVPQTSAYNINSSPTFK-SLSRYESMLRTASSPDSLSRSTPMITPDLE	1626
Qy	1647	SGVQWHLVKNHIEHGDQEGDRGSKMWSEIYTLRLATKGTLOKFVDDLPETIFSTAHRG	1706
Db	1627	SGTKLWHLVKNHDLQREGDRGSKMWSEIYTLRLATKGTLOKFVDDLPETIFSTAHRG	1686
Qy	1707	SALPLAIKYMFDLDEQADKHGHDPVHTWKSNCPLPFRFWNMKNPQVFDIHKNSI	1766
Db	1687	SALPLAIKYMFDLDEQADKHQHDADVHTWKSNCPLPFRFWNMKNPQVFDIHKNSI	1746
Qy	1767	TDACLSVV 1774	
Db	1747	TDACLSVV 1754	
RESULT 7			
Q9UIW1		PRELIMINARY;	PRT; 1328 AA.
AC Q9UIW1			
DT 01-MAY-2000	(TREMBLrel. 13, Created)		
DT 01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT 01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DE OCT/p1exin-A2 protein (Fragment).			
GN OCT/P1EXIN-A2.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=SKLELTAL MUSCLE;			
RC MEDLINE=96149362; PubMed=8570614;			
RA Maestrini E., Tamagnone L., Longati P., Cremona O., Gullisano M.,			
RA Bione S., Tamadini F., Neel B.G., Toniolo D., Comoglio P.M.;			
RT "A family of transmembrane proteins with homology to the MET-hepatocyte			
RT growth factor receptor."			
RL Proc. Natl. Acad. Sci. U.S.A. 93:674-678 (1996).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC TISSUE=SKLELTAL MUSCLE;			
RX MEDLINE=99449305; PubMed=10520995;			
RA Tamagnone L., Artigiani S., Chen H., He Z., Ming G.L., Song H.L.,			
RA Chedotal A., Winberg M.L., Goodman C.S., Poo M.M., Tessier-Lavigne M.,			
RA Comoglio P.M.;			

"Plexins are a large family of receptors for transmembrane, secreted and GPI-anchored semaphorins in vertebrates."

RT EMBL; X87831; CAB57275.1; --  
 DR InterPro; IPR002909; IPT\_TIG.  
 DR InterPro; IPR003659; Plexin-like.  
 DR InterPro; IPR002165; Plexin\_repeat.  
 DR InterPro; IPR001627; Sema.  
 DR Pfam; PF01437; PSI; 3.  
 DR Pfam; PF01403; Sema; 1.  
 DR Pfam; PF01833; TIG; 4.  
 DR SMART; SM00423; IPT; 4.  
 DR SMART; SM00423; PSI; 3.  
 FT NON TER 1  
 FT NON TER 1328  
 SQ SEQUENCE 1328 AA; 147724 MW; 434CF0F0FA42D8D9 CRC64;

Query Match 49.9%; Score 4986; DB 4; Length 1328;  
 Best Local Similarity 70.3%; Pred. No. 0;  
 Matches 935; Conservative 152; Mismatches 239; Indels 4; Gaps 4;

QY 447 FVGTSGKGLKTRVDPGRGNALQYETVQVV-DPGPVLRDMAPSKDHEQLYIMSERQLTRV 505  
 DB 1 FVGTSGKGLKTRADGPPHGGVQYVMSVLKDGSPILRDMAFSIDQRYLYVMSERQVTRV 60

QY 506 PVESCGQYQSCGECCLSGDPHCGWCVLNHTCTRKERCERSKEPRRPAASEMKOCVRLTVHP 565  
 DB 61 PVESCEQYTTCCGLSSGDPHCGWALHNCRRDKCQAWEPNFAAISQCVSLAVHP 120

QY 566 NNISVQNVLLVLETYNVPELSAGVNCFTBELSEMDGLVGNQICQYSPAAKEVPRIT 625  
 DB 121 SSISVSEHSRLLSLVSDAPDLISAGIACAFGNPTVEGQVSGSVQVICISPGPKDVP-VIP 179

QY 626 ENGDDHVVQLOLKSKETGTMTASTSFVFNCSVHNSCLSVSPYRCHCKYRHVCTHP 685  
 DB 180 LQDDWFLGELQURSKETGKIFVSTEFKFNCSAHOQCLSVNSAFRCHCKYRNLCCTHP 239

QY 686 KTCSPQEGRVKLPEDCPQLLRDKLPVPEVVIKPTLKAKNLPQOSGQGVCEILNIQ 745  
 DB 240 TTCSPQEGRIINSEDCPOLVPTTEILIPVGEVKPTLKARNLPQOSGQGVCEVLNIQ 299

QY 746 SEQRPALRFNSSVQCVNTSYSGMEINNLVPELTVVWNGHFNIDNPAQNVKHLKCG 805  
 DB 300 ATRHPALRFNSSVQCVNTSYSGMEINNLVPELTVVWNGHFNIDNPAQNVKHLKCG 359

QY 806 AMRESCGLCKADDPACGCGOCTLRQCHPAQESOWLELSGAKSKCTNPRITFIP 865  
 DB 360 AQRESCGLCKADRPCEGCGERRCTLRQCHPAQESOWLELSGAKSKCTNPRITFIP 419

QY 866 VTGPRGGTKVTRIGENLGLFRDIAHVKGVCSPVLDVGIYPAEQIVCEMGEAKPSQ 925  
 DB 420 VSGPPEGTRVTHGVNGLDFEIAHVVQVAGVCTPLPGEYIAEQIVCEMGEALVGT 479

QY 926 HAGFVEICVAVCRPFMARSSQLYPMTLTLDLKPSPGMSGGTQVTTGNLNAAGSNV 985  
 DB 480 TSGPVRLCIGECKPEFMTKSHQYTFVNPVSLNIRGPESGGTMTITGHYLGASSV 539

QY 986 VMFGKQCLFRRSPSYIVC-NTSSDEVLEMKVSQVDRAKIHODLVPOIVEDPTIVR 1044  
 DB 540 AYLGNQTCFEGRSMSETVVCVSPSSNGLPGVPVSVSDRAHVDNSLQFEYIDDPFRVR 599

QY 1045 IPEPWSIVSGNTPIAWGTHLDLQNPQLRAKHGGKEHINI CEVLNATEMTCPAALALG 1104  
 DB 600 IPEPWSIAGHTPLITGNLNDVQEPRIKRVENGKESVNVKVNVTTLTCLAPSLTDD 659

QY 1105 PDHQSDLTERRPEPGFILDNVQSLILKNTFTYYPNPVFEAPGSPGILKPGTPIILK 1164  
 DB 660 YRPLDTRPDEFGFVFNQVSLIYNDTKFIYENPTFELLSPGVLDPQKPGSPILK 719

QY 1165 GKNLIPVAGNVKLYTVLVEKPCCTVTVSDVOLLCEPNLIGRHKVAKVAGMEYSPG 1224  
 DB 720 GKNLCPAPSGG-AKLYTVLIGETPCAVTVSETQLLCEPNLTGQHKVMVHVGGMVSPG 778

QY 1225 MVIYADPSPSLPAIVSIYAVAGLLIIFIVAVLIAYKRSRESLTLTKRLQWQMDNLESR 1284  
 DB 779 SVSVISDSLLTLPVISAAGSLLIIVILVIAAYKRSRENDLTLTKRLQWQMDNLESR 838

QY 1285 VALECKEAFABLOTDHELITSDLDGAGIPFLDYRTYRWVLPFGIEDHPVLRDLEVPGR 1344  
 DB 839 VALECKEAFABLOTDNELTSDLDGSGIPFLDYRTYRWVLPFGIEDHPVLRDLEVPGR 898

QY 1345 QERVEKGLKFAQLINNKVFLLSFIRTLSEORSFMRDGNVASLIMTVLQSKLEYATDV 1404  
 DB 899 QHVEKALKFAQLINNKVFLLSFIRTLSEORSFMRDGNVASLIMTVLQSKLEYATDV 958

QY 1405 LKOLLADLIDKNLESKNHPKLLLRRTESVAEKMLTNFTFLYKFLKECAGEPLFLPCA 1464  
 DB 959 LKOLLADLIDKNLENKHPKLLLRRTESVAEKMLTNFTFLYKFLKECAGEPLFLPCA 1018

QY 1465 IKOMEKGPIDAITGEARYSLSEDKLIROQIDYKTVLSCVSPDNANSPEVPVKILNCDT 1524  
 DB 1019 IKOMEKGPIDAITGEARYSLSEDKLIROQIDYKTVLSCVSPDNANSPEVPVKILNCDT 1078

QY 1525 ITQVKEKILDAIPKNCVPCSHRPAADMDLEWRQSGARMILQDEDTITTKIENDWKRLNTL 1584  
 DB 1079 ITQVKEKILDAIPKNCVPCSHRPAADMDLEWRQSGARMILQDEDTITTKIENDWKRLNTL 1138

QY 1585 AHYQVDPGSYVALSVKQVTAYNAVNNSTVSRTSASKYENMIRYTGSPDSLSRTPMITPD 1644  
 DB 1139 MHYQVSDRSYVALVPKQTSYNNIPASASISRTSISRYDSSFRYTGSPDSLSRTPMITPD 1198

QY 1645 LESGVKWHLVKQHEHGDQXEGDRGSKWSEIYTRLLATKGTLOKQFVDDLFETIFSTAH 1704  
 DB 1199 LESGVKWHLVKQHEHGDQXEGDRGSKWSEIYTRLLATKGTLOKQFVDDLFETIFSTAH 1258

QY 1705 RGSALPLAIKMYMDELDQADKHGHPHVRHTWKSNCPLRFPWNMIKNPQFVFDIHN 1764  
 DB 1259 RGSALPLAIKMYMDELDQADKHGHPHVRHTWKSNCPLRFPWNMIKNPQFVFDIHN 1318

QY 1765 SITDACLSVV 1774  
 DB 1319 SITDACLSVV 1328

RESULT 8  
 Q9V491  
 ID Q9V491 PRELIMINARY; PRT; 1945 AA.  
 AC Q9V491;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE PLEXA protein.  
 GN PLEXA OR QG11081.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Morten J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Borchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,





1088 VLNATEMTCOPALALGPDHOSDLTERP--EEFGFILDN---VQSLLILNKTNFTYYPNP 1142  
1122 VQSDTECECASPVDV--DSHVEAERPIILEVGFMDNVLVQNLKSVNNHPELYPNP 1179  
1143 VFPAFGSGILELKPPTIILKGNLIPPVAGGNVKNLYTVLGEKPCCTVT--VSDVOLL 1201  
1180 EYIFEEER--VKYFKSYLTINGRNLDRACKESDVE---VKIGNGFCNITSLSRQQLTC 1233  
1202 ESN-----LGRHKMARVG--GMEYSPGMV--YIAPDSPLSLPAIVSIAVAGLLIIFI 1253  
1234 RPPSEATATKSMNGPEVIVRIGTSLEYRIGLSYESNNIILDMGENVIFAVIATIVILL 1293  
1254 --VAVLIAYKRRSRLTLKRLQOMDNLESVALECKEAFELQTDIELHETSDLDGAG 1311  
1294 IFVALLVAYKKSESERVLNRQEQMDILELRVAECKEAFELQTDIELHETSDLDGAG 1353  
1312 IPFLDYRTYMRVLPFGIEDHPVLRDLVPGYRQ--RVREKGLKFAQLINNKVFLLSFI 1369  
1354 IPFLDYRSYAMKILFPNHEDHVLQ-----WERPELLRKEKGLRIFQGLIMNKTFLLFI 1408  
1370 RTLESQSFMRDGNVASLIMTVLQSKLEVATDVAKQLADLIDKNLESKNHKKLILRR 1429  
1409 RTLESNRYFSMRERVNVASLIMTVLQSKLEYCTDILKTLGLDLIEKIEGKSHPKLLRR 1468  
1430 TESVAERKMLTNWTFLLYKFKCECAGPLFSLFCAIKQOQMEKGPIDAITGEARYSLSEDK 1489  
1469 TESVAEKMLSANWTFLLYKFKCECAGPLYMLFRAVKGQVDKGPVDACCTHEARYSLSEK 1528  
1490 LIRQIDYKTL-----VLSVSPD--NANSPEVPVKILNCDTITQVKEKILDAIFK 1538  
1529 LIRQIDFRPMNVYASIIQOPIFCNLDMLPSHTENVSVKVLDCDTITQVKEKILETIYR 1588  
1539 NVPCSHRPKADMDLEWRQSGARMILQEDITTKIENDWKRLNLAHYQVDPGVSVALV 1598  
1589 NIPSPRPKDDLDLEWRTGATGVILYDEVDVTSKESWKMLNQLQHYNVPDGCAGSLV 1648  
1599 SKQVATYANVNNSTVSRTSASKYE--NMIRYT--GSPDLSRTRTMITPDLESYGVMHVLV 1655  
1649 PKQSSIYFSLSDKNEKS--HKYETLNIISKYTSSTPFSRAGSPLNMDHENGRLYHVLV 1707  
1656 KNHEHGQKEDGRGSKVMSEIYLRLLATKGTLOKQFVDDLFETTFSTAHRSALPLAIKY 1715  
1708 KHDSDMQKEGERVKNLYSEIYLRLLATKGTLOKQFVDDLFETTFSTAHRSALPLAIKY 1767  
1716 MFDFLDQADKHGHDHVRHTWKSNCILPFRWNMIKNPOFVEDIHKNSITDACLSSVA 1775  
1768 MFDFLDQALHGHTDPEVHTWKSNSLPLFRWNLIKNPNFVFDIHKNSIVDSCLSVA 1827  
1776 QTFMDSCTSHRLGKDSPSNKLAYKADIPSYKNWVERYYSDIGKMPAISDQDMNAYLAE 1835  
1828 QTFMDSCTSDHRLGKDSPPSKLAYKADIPYRKWVDYRDYRDMSPISDQDMNAYLAE 1887  
1836 QSRMHMEFNTMSALSEIFSVGVKSEIILGPDHDDQCGKQKLAYKLEOVITLMS 1891  
1888 ESRLLHTEFNTNCALHELITYAVKYNEQLTVLSEDEFQKRLAFKLEQVHNIMS 1943

RESULT 10  
Q96GN9 PRELIMINARY; PRT; 813 AA.  
ID Q96GN9  
AC Q96GN9

Q96GN9;  
01-DEC-2001 (Tremblrel. 19, Created)  
01-DEC-2001 (Tremblrel. 19, Last sequence update)  
01-MAR-2002 (Tremblrel. 20, Last annotation update)  
Unknown (Protein for IMAGE:4130636) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;

Strausberg R.;  
Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC009343; AA09343.1; -  
DR InterPro; IPR002909; IPT\_TIG.  
DR Pfam; PF01833; TIG; 1  
FT NON\_TER  
SQ SEQUENCE 813 AA; 91889 MW; A310511C01E2A237 CRC64;  
Query Match 33.9%; Score 3386.5; DB 4; Length 813;  
Best Local Similarity 79.1%; Pred. No. 5.2e-235;  
Matches 641; Conservative 76; Mismatches 92; Indels 1; Gaps 1;  
Qy 1085 ICEVLNATEMTCOPALALGPDHOSDLTEREERFGFILDNQSLLILNKTNFTYYPNPVF 1144  
Db 5 VCKVNTTTLTCLAPSLTTRPGLDITVERDEFGFVFNVSQSLIYNDTKFIYYPNPTF 64  
Qy 1145 EAFGSGILELKPCTPIILKGNLIPPVAGGNVKNLYTVLGEKPCCTVTVSDVOLLCESP 1204  
Db 65 ELLSPTGVLDQKPSPIILKGNLCPASGG--AKLNYTVLIGETFCATVTSQTLLCEPF 123  
Qy 1205 NLIQRHKVMARVGMEXSPGMVYIAPDSPLPAIVSIAVAGLLIIFIVAVLIAYKRS 1264  
Db 124 NLTGQHKVMHVGMVSPGVSIVSDLSLLTLPAIVSIAAGGSLLLIIVILVLIAYKRS 183  
Qy 1265 RESLTLKRLQOMDNLESVALECKEAFELQTDIELHETSDLDGAGTFLDYRTYTMV 1324  
Db 184 RENDTLKRLQOMDNLESVALECKEAFELQTDIELHETSDLDGAGTFLDYRTYTMV 243  
Qy 1325 LFPGIEDHPVLRDLVPGYRQERVEKGLKFAQLINNKVFLISFIRTSQSFMRDRG 1384  
Db 244 LFPGIEDHPVLRDLVPGYRQERVEKGLKFAQLINNKVFLISFIRTSQSFMRDRG 303  
Qy 1385 NVASLIMTVLQSKLEYATDVILKQLADLIDKNLESKNHKKLILRRTESVAERKMLTNWTF 1444  
Db 304 NVASLIMTVLQSKLEYATDVILKQLADLIDKNLESKNHKKLILRRTESVAERKMLTNWTF 363  
Qy 1445 LLYKELCECAGPLFSLFCAIKQOQMEKGPIDAITGEARYSLSEDKLIRQIDYKTLVLS 1504  
Db 364 LLHKFLCECAGPLFSLFCAIKQOQMEKGPIDAITGEARYSLSEDKLIRQIDYKTLVLS 423  
Qy 1505 VSPNANSPEVPVKILNCDTITQVKEKILDAIVYKNVPSQRPRAVMDMLEWRQGIARVV 1564  
Db 424 VSPNANSPEVPVKILNCDTITQVKEKILDAIVYKNVPSQRPRAVMDMLEWRQGIARVV 483  
Qy 1565 LQEDITTKIENDWKRLNLAHYQVDPGVSVALVSKQVATYANVNNSTVSRTSASKYENM 1624  
Db 484 LQEDITTKIENDWKRLNLAHYQVDPGVSVALVSKQVATYANVNNSTVSRTSASKYENM 543  
Qy 1625 IRYTGSPLSRTRTMITPDLESYGVMHVLVKNHEGDKQEGDGRGSKVMSEIYLRLLAT 1684  
Db 544 IRYTGSPLSRTRTMITPDLESYGVMHVLVKNHEGDKQEGDGRGSKVMSEIYLRLLAT 603  
Qy 1685 KGTLOKQFVDDLFETTFSTAHRSALPLAIKYNMDFLDEQADKHGHDHVRHTWKSNCIL 1744  
Db 604 KGTLOKQFVDDLFETTFSTAHRSALPLAIKYNMDFLDEQADKHGHDHVRHTWKSNCIL 663  
Qy 1745 LRFVWNMIKNPOFVEDIHKNSITDACLSSVAQTFMDSCTSHRLGKDSPSNKLAYKADI 1804  
Db 664 LRFVWNMIKNPOFVEDIHKNSITDACLSSVAQTFMDSCTSHRLGKDSPSNKLAYKADI 723  
Qy 1805 PSYKNWVERYYSDIGKMPAISDQDMNAYLAEQSRMHMEFNTMSALSEIFSVGVKSEI 1864  
Db 724 PSYKNWVERYYSDIGKMPAISDQDMNAYLAEQSRMHMEFNTMSALSEIFSVGVKSEI 783  
Qy 1865 LQPLDHDQCGKQKLAYKLEOVITLMSLDS 1894  
Db 784 LQPLDHDQCGKQKLAYKLEOVITLMSLDS 813  
PRELIMINARY; PRT; 2051 AA.  
ID Q96682  
AC Q96682;  
RESULT 11  
Q96682  
AC Q96682

DT	01-MAY-1999 (Tremblrel. 10, Created)	Db	584	QOCIEFESIIPEKIPISLSELHLLIIRT--LPEPNAKRYCFVGNSTPIDAEILENGLGC	641
DT	01-MAY-1999 (Tremblrel. 10, Last sequence update)	Qy	613	YSPAKEVPRITITENGHDVVVQLQKSKETGWTASTSFVYNGSVHNSCLSVESPVRC	672
DE	01-JUN-2002 (Tremblrel. 21, Last annotation update)	Db	642	ATPPLDERP-LIPTNDHILVPLSVRSSETNKDFVSRNFAFFDCSHGNCOCESQSSWGC	700
GN	PLEXB OR CGI7245.	Qy	673	HWCKYRHVCTHDPKTCFQSGVRKLPEDCPQLLR-VDKILVPVEVIRPITLAKNLPQPOQ	731
OC	Drosophila melanogaster (Fruit fly).	Db	701	NWCFIDNKCVKHSLOCRNIENAVSTVGHCPHLKKNRPEILLPVRVPTEIRLEIENLPKPK	760
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	Qy	732	SGORGECILNIQGSSEORVPALRPNSSVQONTSYSEGMENINLPELVTVVWNGHNI	791
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	Db	761	SAHAGFLCTIIEAQQMLLPAHVESNNKVVXCTPYFYE-XNTHYQAKVVTWNFOHYV	819
OC	Ephyrdoidea; Drosophilidae; Drosophila.	Qy	792	DNPAQNKVHYKC--GAMRE--SCGLCKADPDPACWCQCGOCTLRHCPAQESQWL	846
OX	NCBI_TaxID=7227;	Db	820	DTAI---VTLKCDVLGSHREHPCSLCVTRDPKYKCAWCS--NSCVNETCIADKNS--	872
RN	SEQUENCE FROM N.A.	Qy	847	ELSGAKS---KCTNPRITEIIPVTPREGGTVKTIIRGENLGLFRDRIASHVKVAGVECS	902
RP	MEDLINE=99091049; PubMed=9875845;	Db	873	ISSGSKSAIENECPLPRIDIILKPLSGPVEGTLITIEGSLGIREEDVRGKIFIGSVPC	932
RA	Winberg M.L., Noordermeer J.N., Tamagnone L., Comoglio P.M.,	Qy	903	PLVDGYIPAEQIVCEMGEA-----KPSOHAGFVILCAVCPPEPMARSQLYFMTL	954
RA	Striggs M.K., Tessier-Lavigne M., Goodman C.S.;	Db	933	--LENYQISVKIECRTGASLYEMSAPIKANDAGFTE-----SSVQPHFKNV	977
RT	"Plexin A is a neuronal semaphorin receptor that controls axon	Qy	955	TLSDLKPSRGPMSGGTQVTTITGNLNAAGSVVWVFGKQPCLFH--RRSPSVIVCNTT---	1009
RT	guidance.";	Db	978	LTGLIPTIGRSGTQSLIGLKGFLNIGSTWRAPFLDEVECHIDVTQASSQSVCTTSAT	1037
RL	Cell 95:903-916(1998).	Qy	1010	-----SSDEVLEMKVSU-----QVDRAKITHODLVFOYVEDP	1040
DR	EMBL; AF106933; AAD09426.1; --	Db	1038	QPEPIRSLHLVIDGANKRTLECSQISTPSITNTNTPRSNFGSYQLRALPRQPCSIENYTD	1097
DR	FlyBase; FBgn0025740; plexB.	Qy	1041	TVIRIPEWSIVSGNTPVAVGTHLDLIQNPQIPRAKGGKEHIN--ICEVLNATEMTQOA	1098
DR	InterPro; IPR002909; IPT TIG.	Db	1098	RIMQIKPLRSFKSGRVLTVHGIYLSIQKPELEVY--DNERVNKTSVCVINSNQMECP	1156
DR	InterPro; IPR003659; Plexin-like.	Qy	1099	PALALGPD-----HOSDLTERPEFG-----	1119
DR	InterPro; IPR002165; Plexin repeat.	Db	1157	PPVNYKPETKNRMKMDTELHQNSSFQETRYKNEYDKRRRADFGDTFLYTNAGSS	1216
DR	Pfam; PF01437; PSI; 3.	Qy	1120	-----FILDVNSQLILNK-----TNFTYYPNPVFEA	1146
DR	Pfam; PF01833; TIG; 3.	Db	1217	ANYFVNNMVDVTFVKVHETQLNLQLSFVMDVQLVRNLNKVYFDIRSTIVVLADPKYL	1276
DR	SMART; SM00429; IPT; 3.	Qy	1147	FGSGILELPGTPPIILKGNLIPPVAGNVKLYTVLVGKPKCTVT--VSDVQLLCESPN	1205
DR	SMART; SM00423; PSI; 3.	Db	1277	FPNDGV-KLYKGDVTVIEGELL--NLAADEYDNNVT--IGTSQCNITSLALNQLLCIPPE	1331
SQ	SEQUENCE 2051 AA; DBF27F9CA8FF7EA CRC64;	Qy	1206	-----LIGRHKVMARVCGMEVSPGMVVIAPDPSLSLPAIVSIA	1243
	Query Match 29.4%; Score 2932.5; DB 5; Length 2051;	Db	1332	QQPLPTDENGVDGSTDPLVVRVGRN-LRPVIGLYKIDLNKPYV--SHALLVGLITVA	1388
	Best Local Similarity 34.4%; Pred. No. 1.3e-201;	Qy	1244	VAGGLLIIFTAVLLIAYKRSRESDLTKRLQMDNLESVALECKEAPAELOTDIHEL	1303
	Matches 708; Conservative 354; Mismatches 682; Indels 315; Gaps 64;	Db	1389	-----LLVWVVFVILIIIPRRKSTQAEYKRIQIOMITLESNVRSCKQAFELQTDMDL	1444
Qy	51 FNHLVVDERTHYIGAVNRIYKLSDDLKVLVYHETGPDENPKYPPRIVQFCNPLTT	Qy	1304	TSLDGAGIFPDLRYTVMRVLPFGIEDHPVLRDLEVPYR---QERVEKGLKLAQILI	1359
Db	87 FTHMSFDFMHNVLFAGATNKILKNLENRLVAEAVTGPLHDSPOCH---AGGCPED	Db	1445	TADLESTGIPTLDHVNIMKVFVFGVSDHTLNS---PKFREGSPQNTYDAAMVQFEQI	1501
Qy	111 T--NNVKNMLLDY-KENPLIACGSLYQICLKRL-----EDLPKLGEPHYKKEHYLS	Qy	1360	NNKVFLLSFTLTLESOR-SFSMRDRGNVASIMTVLOSKEVATDVLKQLADLIDKLE	1418
Db	143 SLVNNFKILVSVYAHGDLIACGSIQACIYSLPRFPATPQFFAVP-----LA	Db	1502	GNKYFLLMFLETUEAQRSSFSIRQVNVASLIMVLMNKMMEYATEILKSLRLRLDKSLA	1561
Qy	162 GYNESGVFGV1--VSYS-NLDDKLFIAIATVADGKPEY---FPTISSRKLTKNSEADG	Qy	1419	SKNPKLILLRTSVAEKMLTNWFTFLLYFLKECAGEPLPSLFCAIKQOMEKGBIDAIT	1478
Db	194 ANDENASTVAFVGPARYANKEEDILVGTFTVNGVYRHDVPAISRRLDNLNAYE	Db	1562	SK-HPQLMLRSTSVSEKMLTNLYAICMYDYLYEYAGSNLFLPKAIKHQIEKGLVDAIT	1620
Qy	216 YVHDFEFVASKIKIPSDTFTIIPDPDIYVYVYFGSSGNFYVFLTQPEWSPPGSTTKEQV	Qy	1479	GEARYSLSEDKLIQQIDYKTLVLSVCVSPNANSPVVPVKILNCDDTTITQVKEKILDAIFK	1538
Db	250 PSIQQSIINIDVKYR-----HFLVDYVYGFNSSEYAFIIVQKK-----SHLAD	Db	1621	NDARYSLSEERLLHEQVTHSVILHILQDD--LDEKVCQRVNDWDTISQVKKLILDAIFK	1678
Qy	276 YTSKLVRLCKEDTAFNSYVEVPGC--ERSGVEYRLLQAYLSKAGAVLGRTLGVHPDD				
Db	298 YVTRLARICITDPNDYSYTEITVQCTATENNVDYNTIRDAKVTPASHKLAQRMGIKKDH				
Qy	334 LPLTFVSKQKRMKSLDESALCIFILOKNDRIKERLOSCYRGEGTLDLAWLKVDIPC				
Db	358 VLVTFVSPREISNQPESSAMCIYSIKIEDFMFIENIHLCFNG-----TTKORNL				
Qy	394 SSALITIDP-----NFCGLDMNAPLGSDMVGRIPVFTEDRDRMTSVIA--Y				
Db	409 GYISGTINDGRCPVIGSLGNIYVFCVGLKIS-GVSP-ITAHALPHFDNVSVTSVATST				
Qy	439 VYKNHSLAFVYKSGKLKIRVDGPRGNALQYETVQVDPG-PVLDMAFSKDHEQLIYM				
Db	467 TDQOHSALFLLGNMVGKLVLLSGQSPG--EYBEI-VVDAGNRILPNTMWSPPKOFLYVL				
Qy	498 SRROLTRVPVSGQYQSGCEGICGSDPHCGWCVLHNTCTRKERCERSKEPRFAS--EM				
Db	524 SORKITKLRIEHCVSNTNCSACLESDDPCGWCSEKRCRTVSTCORDTSASRWLSLGS				
Qy	556 KQCVRL-TVHPNNISVQYNVL-LVLETYNVPE-LSAGVNCVTFEDLSEMDGLVVGNIQC				



Db 1305 DTIEMKATEKEFAFAELQSLNQYADLPGLTAPFLEYKDYCARVLPNAGKNAVL 1364  
Qy 1336 RDLVPGVROERKEGKGLPAQLNNKVFLLSIRLTLESORSFMRDRGNVASLITVLQ 1395  
Db 1365 KNLVETHKAEALTEAGREFFHKLMMKNTFLUTWVRIMEANKYFVGKDRVYVGSLLMVVLQ 1424  
Qy 1396 SKLEYATDVLKQLADLIDKNLESKHPKLLLRRTESVAEKMLTNWFTFLLYKFLKEC-A 1454  
Db 1425 EXMGVCTEMLKQLRELIEKTEKQPKILFRSESESIAERMLAAWFTFLHDLKTYDA 1484  
Qy 1455 GPFLESLFCAIQOMKEGPDIDAITGARYSLSDKLIROQIDYKTLVLSVSPDN--ANS 1512  
Db 1485 GKLYELFWGIKQOMKEGPDALTLEARYSLSEKILRATFEYKELMVVFVASDSVYSST 1544  
Qy 1513 PEVPVKILNCDDITTOYKEKILDAI FKNVPCSHRPAKADMDLEWRQGSARMILQDEITT 1572  
Db 1545 QDIPVRLVDCDDITTOYKEKILDAI FKNVPCSHRPAKADMDLEWRQGSARMILQDEITT 1604  
Qy 1573 KIE-NDMKRLANTLAHYQVDPGSGVVALVSQVTAYN-AVNNTSVTSASKEYENMI---RY 1627  
Db 1605 RVEGGNWKELNTLAHYVNPNAITLTSKNSLYNLSILSDRSEKSSVMKTPGVGSPK 1664  
Qy 1628 TCSPLSRLSRTPMITPDLESQVQWHLVKNHEGDOKEGDRGKMVSEIYLTLLATYKT 1687  
Db 1665 WGAPPEVIAADHASSADTDNGFKLYHLVKPTEHGPT---DSQEKMVTEIYLTLLAMKGT 1721  
Qy 1688 LQKFVDDLPETTFSTAHRSALPLAIKYMPDFLDEQADKHGIDHPHVRHTWKNCLPLRF 1747  
Db 1722 LAKFITSLESIFSS---STVPPCIKMPDFMDEQAREHGINDEPVVHAWKSNALPLRF 1777  
Qy 1748 WYNNIKNPOFVFDIHKNSITDACLSVVAQTFMDSCSTSEHRLGKDSPSNKLKYAKDIPSY 1807  
Db 1778 WYNNIKNPHFDIQKPTKIEGCLSVVAQTLMDACSTQDHLTKDSPSKLLFAKMYQY 1837  
Qy 1808 KMWERYYSIDIGOMPAISODMNAIYLAOSRMENFNTWSALSEIFSVMGKYSIEILGP 1867  
Db 1838 RDLVDSYYTEISMTPTRIDDATMGANLSAESRQHNGEFHVSALNELYKYLDQKESIIDA 1897  
Qy 1868 LDHDDCGQCKLAYKLEQVITLMSLD 1893  
Db 1898 LESNEHAQASRLPGRLODLALMESD 1923

## RESULT 13

Q9N375  
ID Q9N375 PRELIMINARY; PRT; 1944 AA.  
AC Q9N375;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Y55F3AL.1 protein.  
GN Y55F3AL.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX Waterston R.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC024825; AAF60788.1; .  
DR InterPro; IPR002909; IPT\_TIG.  
DR InterPro; IPR003659; Plexin-like.

DR InterPro; IPR002165; Plexin\_repeat.  
DR InterPro; IPR001627; Sema.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF01437; PSI; 2.  
DR Pfam; PF01403; Sema; 1.  
DR Pfam; PF01833; TIG; 4.  
DR SMART; SM00429; IPT; 3.  
DR SMART; SM00423; PSI; 3.  
DR PROSITE; PS00284; SERPIN; UNKNOWN 1.  
SQ SEQUENCE 1944 AA; 216978 MW; D6C5EF7DEA2BD8A8 CRC64;  
  
Query Match 28.4%; Score 2841; DB 5; Length 1944;  
Best Local Similarity 34.3%; Pred. NO. 4.7e-195;  
Matches 674; Conservative 361; Mismatches 733; Indels 198; Gaps 54;  
  
Qy 49 EGFNHLVVDERTGHIYLGAVNRIYKL--SSDLKVLVTHETGDEDNPKC----- 95  
Db 41 EVFEKMAIDPSTTRFVFGAVNSYDLTSDATVRRHVQTGPDSDSPLCRGKWNKNSLKK 100  
Qy 96 -YP-----PRIVOTCNEPLTTNNNNKMLLDYKENRLIACGSLYQICLKLRLEDLFLGL 150  
Db 101 LYQOKNVARNRRECBHQLSRNHSHTKALSVDKSKLIECSNLYQGR-----RLRLNHNIS 157  
Qy 151 EPHKK-EHYLSGVNSESFGVIVSYNSLDDK--LFIATAVDGKPEYFPTISSRLTKN 207  
Db 158 EVISEALEPRVSNDSSTSSVIFVCGSPANLSSEPVLVYGA-----TIGSGD--- 203  
Qy 208 SEADGMFAYVPHDEF-VASMIKIPSDFTIIT-----PFDIYVYVG 247  
Db 204 -----HDMSSVSLFLRPQKAFEVFPGLYGTHVSLDYSRGGYKYQIDYING 252  
Qy 248 FSSGNFVFLTLQPEMVSPPGSTTKQVYVTSKLVRLCKEDTAFNSYVEVPICERSGVBY 307  
Db 253 FESGDYAYFTRQRVNADSS-----IQSLRVVCTGDKNFHSHYTEVPLECTONGVEF 306  
Qy 308 RLQAAVLSKAGAVLGRTLGVHPDLDLFTVFSKGQKMKSLD---ESALCIFILKQIN 364  
Db 307 NLQDVVTVTRAGYELAKSLDISVSDPVLYGVFVWEGDKNSYRSEPTGKSAICMFTWREIE 366  
Qy 365 DRIKERLQSCYRGEGTL--DLAWLKVKVDIPCSSALLTIDDNFCGLDMNAPLGVSDMVRGI 422  
Db 367 TSFKQNTMKCYGTSGLKKNLPWFSSND-DCRFTTLPWEGIKCGKGVNSKIGDTPISTS 425  
Qy 423 PVFT-EDRDRMTSVIAY-VYKNHSLAFVGTSGKLLKIRVDGPRGNALQYETVQVVDGP 480  
Db 426 ATYVWEDSSNLLTAINTTSSSTVAFVGTGGQLHLILIESKR-SAEKYATEMLTDNEP 484  
Qy 481 VLDMAFSKDHEQLYIMSRQLTRVPVSECGQYOS-CGECILSGDPPHCGVCLHNTCTRK 539  
Db 485 ILSDMFEGDGKHYVILTPSKVIKMPTRSCEALSTQCDTCLSSRDYPCGWCVSNHCTQE 544  
Qy 540 ERCERSKEPRRPAS-----EMKQCVRL-TVHPNNTSVSQYVNL-LVLEYVNPVPELSAGVNC 593  
Db 545 ESCER-EVPHTSRGWLDFQNSKCPRIIRSVKPDQIQINTADYLVNTIENLOAPK-GRMOC 602  
Qy 594 TFE-----DLSEMDGLVGNQIQYCSAPAAKVPRIITENGDDHVVQLOLQKSETGTMTFAST 649  
Db 603 LQFSTGDAVVDPMFPGDSLKCATPPMNRILPRIPT-NEYHLPAKLIIVSDGSKLPLATT 661  
Qy 650 SFVPYNSVHNSCLSVESPYRCHWKYRHVCTHDPKT-----CSFQE 692  
Db 662 NFSFYDCNRYTSCSTCSASQPPCDWCLNESCEVAGKLTEDCKRQKHVGNLNRDGSIRK 721  
Qy 693 GRVKLPEDCPQLLR-VDKILVPEVIEKPIITLAKNLPQPSQGRGYECILNIQSGEQRVP 751  
Db 722 G-----PSKCPHIVAPVSKMSVATGERNNISVKVENVDATFMGD--FKCFKYGTVTHEKI 775  
Qy 752 ALRPNSSVQONTSYSEVGEI---NNLPVELTVWVNG-----HFNIDNPAQWK 798  
Db 776 AMRTSDDTITCDEMFLPEFYGTSLGSGSTPGVFNVINSATGFAVGKKSINRVLNVDVASLA 835  
Qy 799 VHLYKCGAMRESGLCLKADPD-PACGWCQPGGCTLHQHCPAQE--SOWLELSGAKSKC 855

Db 836 IDVSCENLATNCRCLTLADADKYDCGWCSDKSCARPHQCNSRQIPENWL--NASQLC 892  
QY 856 TNPRITEIIPVTGREGGKTIVIRGENLGLFRDIASHKVAGVECSPLVDGYIPAEQIV 915  
Db 893 PNPVIEDFKPKGSGIFGGTRVINGINLGRHVSDEKAVQIANVPCB--VVEYVPSQKIV 950  
QY 916 CEMGBA--RPSQAHAGFVEICVACRPEFMASSQLYPMFTLTSLDKPSRPMGSGTQVT 973  
Db 951 CVTGKSPKSGNERGVAVTLRHSCLKFFAHSTEHFSYQPVVSSMKPVKSPRSRGTQVT 1010  
QY 974 ITGTNLNAGSNVVMFGKOPCLPHRRSSYIVCNTTSSDEVLEMKYSVQVD--RAKIHQD 1031  
Db 1011 LTGVLDLTGAETVRFQGTGCKILERASGLWRCRMGASGEGQFPLHIAFGQLQNLPIP 1070  
QY 1032 LVFQVVEDPTIVRIPEWSIVSGNTPIAVWGTDLHLQNPQIRAKHGG--KEHINICEVL 1089  
Db 1071 IYFETSDPSVQISPNKTIIPSGGIVDVYGGFTLLQRPVMTGLGDEKSYGPACSIE 1130  
QY 1090 NATEMTCOAPALAGPDHQSOLDTERPEBFGFILDNVOSLLILNKTNFTYYNPVFEAFGP 1149  
Db 1131 DDQMKCATPT-APGKSDRSGQSERITEYAFDFDGAEE--LYRSRIKVMPSPTPETLSE 1185  
QY 1150 SGILELKPGTPIILKGNLIPPVAGNVKLNVTYVLGKPKCTVT-VSDVQLLCESPNLIG 1208  
Db 1186 PRF--VRPGEDFTLNGHNLNLAAS--ERDIDVKVGGEPCPLTALANKVLTQPPVKKP 1240  
QY 1209 R-----HKVMARVCGMEYSGPMV-YIAPDSPLSLPAIVSIAVAGGLIIIFIVAVLIAYK 1261  
Db 1241 RGAGKLNPEIIVTVGNASISIGEVSYSGLSSSLFFIIMFTII--FMALIIICLIVLYR 1298  
QY 1262 RKSRESDLTKRLQOMQDNLESVALECKEAFAELOTDIHELTSD--LDGAGIPFLUDYRT 1319  
Db 1299 RKSINTHOROMKYLKTMQDTEMKVATECKEAFAELOTLNQVYADTLPLGTPTAPFLEYKD 1358  
QY 1320 YTMRVLPFGI-EDHVEVLRLDLEVPYQROERVEKGLKPLAQLINNKVFLLSFRTLESQRSF 1378  
Db 1359 YCARVLFPNAGKNHVALKNLEVDTHKAEAEAGLREFHKLNMNKTFLTMTVKTMEANKYF 1418  
QY 1379 SMRDGNVAGSLIMTQLSKLEYATDVLLKOLLADIDKNLESKNHPKLLLRRTESVAEKML 1438  
Db 1419 VGKDRVYVGSLLMVVLOEKMGYCTEMKQLLRELIEKTEVKEKFPQKILFRSESIAERML 1478  
QY 1439 TNWFTFLLYKPKEC-AGEPLFSLFCAIKQOMEKGPIDAITGEARYSISEDKLIROOIIDY 1497  
Db 1479 AAWFTLLHDLKTYDAGKKLYELFWGIKQOMEKGPQDALTLEARYSLSEBKLLRATFEY 1538  
QY 1498 KTLVLSCVSDN--ANSPEVPVKILNCDFITQVKEKILDAIFKNVPCSHRPKAADMDLEW 1555  
Db 1539 KELMVFSASDSVYSSTQDIPVRVLDCDFTQVKEKCLDARKYGRFADRPAGDMELEW 1598  
QY 1556 RQSGARMILQDEDTTKIE-NDWKRLNTLAHYQVPGSVVALVSKQVTAYNAVNNSTVS 1614  
Db 1599 KTGMMKMALQDIDSSRVEGGNWKRLNTLAHYNVPNNAILTTSKNSLYNLVNFSSTS 1658  
QY 1615 RTSASKYENMIRYT-----GSPDSLSRSPMTIPDLESYGVKMWHLVKNHHE 1659  
Db 1659 MKTASILSDRSEKSSVMKTPGVGSPKPGWAPPVIAHASSADTDNGFKLYHLVKPTE 1718  
QY 1660 HCDQEGDRGSKWSEIYLTRLLATKGTLOKFVDDLFETIESTAHRGALSALPLAKYMFDF 1719  
Db 1719 HGPT---DSQBKMVTIEYLTRLLMMKGTTLAKFITSLLESIFSS---STVPPCIKTMFDF 1771  
QY 1720 LDEQADKHGIDHPVHRHTWKSNCPLPLRFWNNMIKNPQFVFDIHKNSITDACLSVVAQTFM 1779  
Db 1772 MDEQAREHGINDPEVVHAWKSNALPLRFWNLNKNPHFLFDIQKPTKIEGCLSVVAQTFM 1831  
QY 1780 DSCSTSEHRLKQDSPSNKLLYAKDIPSYKNWERYYSIDGKMPAISDQDMNAVLAEOGRM 1839  
Db 1832 DACSTQDQHTKQDSPSKLLFAKDMYQYRDLVDSYYTEISMTPRIDDATMGANLSAESRQ 1891  
QY 1840 HNNEPNTMSALSEIFSYGVKYSSEITGLDHDHDDQCGKOKLAYKLE 1884  
Db 1892 HNGEFHVFSAUNELYKLDQYKESIDALESNEHAQASRLPGRLO 1936

## RESULT 14

Q9UJ92 PRELIMINARY; PRT; 1952 AA.  
AC Q9UJ92;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE PLEXIN-B1/SEP receptor precursor.  
GN PLEXIN-B1/SEP.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99449305; PubMed=10520995;  
RA Tamagnone L., Artigiani S., Chen H., He Z., Ming G.L., Song H.L.,  
RA Chedotal A., Winberg M.L., Goodman C.S., Poo M.M., Tessier-Lavigne  
RA Comoglio P.M.;  
RT "Plexins are a large family of receptors for transmembrane, secreted  
and GPI-anchored semaphorins in vertebrates.";  
RL Cell 99:71-80 (1999).  
DR EMBL; AJ011415; CAB56222.1; --  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR002309; IPT\_TIG.  
DR InterPro; IPR003566; Lipocln\_cytFABP.  
DR InterPro; IPR003659; plexin-like.  
DR InterPro; IPR002165; plexin\_repeat.  
DR InterPro; IPR001627; Sema.  
DR Pfam; PF01437; PSI; 2.  
DR Pfam; PF01403; Sema; 1.  
DR Pfam; PF01833; TIG; 3.  
DR SMART; SM00429; IPT; 3.  
DR SMART; SM00423; PSI; 3.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_1.  
KW Receptor; Signal.  
FT SIGNAL 1 30  
FT CHAIN 31 1952  
SQ SEQUENCE 1952 AA; 214410 MW; FCAD0630E128EE9C CRC64;

Query Match 28.0%; Score 2792.5; DB 4; Length 1952;  
Best Local Similarity 33.2%; Pred. No. 1.5e-191;  
Matches 687; Conservative 353; Mismatches 704; Indels 325; Gap

QY 12 LSHLLMVGSGSSTLLTRQAPLPSQKORSFV---TFRGEPAEGFNHLVVDERTGHIYLG  
Db 4 LGPALLQALWAGVWLTLPPLP-----PTAFTPNGTY-----LQHLARDPTSGTLYLG  
QY 69 NRIYKLSDLKVLVTHETGPDENPKCVPRIVQTCNEPLTTNNVNMKLLIDYKBN  
Db 53 NFLFQLSPGLQLEATVSTGVLDSRDCLPPVMPDECPQ-AQPTNNFNQLLV--SPGA  
QY 129 ACGSLYQGI CKLRLLEDLFLKL-----GEPYHKHEHYLSGVNESGSVFGVIVSYNLL  
Db 110 VCGSVHQVCEQRRLGQLEQLLRPERPGD-----TOYVAANDPAVSTVGL-----  
QY 182 KLFTATAVDGKPEYF-----PTISSRKLTKNSEADGMFAFYVPHDFEYVAS  
Db 156 ---VAQGLAGEPLLVFVRGYTSRGVGGGIPPIITRALWP-PDPOAIFY-----EETAK  
QY 228 KIPSDTFTIIPDFDIYVYVFGSSGNFVYFLTLQPEMVPSPGTTKEQVYTSKLVRLCK  
Db 207 -----AVGRLSEYSHHFVSFAFAGASAYFLFLRRDL-----QAQSKAFRAYVSRVCL  
QY 288 TAFNSYVEVPIGCRSGVEYRLLQAAKYLKAGAVLGRTLGVHPDDDLLFTVPS-----  
Db 256 QHYYSYVELPLACE--GGRYGLIQAQAAVATSRV-----AHGEVYFAAFSSAAPF  
QY 341 -KGQKRKMKSLDESALCIFIILKQINDRIKERLQ-SCY-----RGEGLTDLWLKVKDIF

Db 306 GRPPSAAGASGASALCAPFLDEV-DRLANRTRDACYTREGRAEDTEVAYIEY-DVNSD 363  
QY 395 SALL---TIDDPFCGLDMN-APLGVSDMVRGIPVFTEDRDMTSTVIAYVYKHSIAFVGT 450  
Db 364 CAQLPVDTLDAFPCSDHTFSPWASRPLEANTPILWPGIQLTAVAVTMEDGHTIAFLGD 423  
QY 451 KSGKLKIRVDGRGNALQYETVQVVDPGFVLDRMAFSDKHDEQLYIMSERQLTRPVVESC 510  
Db 424 SQQLHRVYL-GFGSDGHPYSQSIQOGSAVSRLTFDGTFFHLYVYMTQSTLLKVPVASC 482  
QY 511 GOYQSGCEGLSGDHPCHGCVLHNTCTRKERCSEKPRR----FASEMKQCVRLL-TVHP 565  
Db 483 AQHLDCASCLAHRDPYCGVLLGRCSSECSGQGPQWLSFQPEL-GCLQVAA MSP 541  
QY 566 NNISVSQAVNLVLVETVNPPEL-----SAGVNTCFEDLSEMDGLVGNQIQCVSPAKEYV 620  
Db 542 ANISREETREVL-----SVFDPPLPWPBGYSYCHGE-HQSPALLTSGVWCMSPDSEA 596  
QY 621 PRIITENGHHVVQLKSKETGMTASTFVFNCSV-----HNSCLSCVESPYRCHW 674  
Db 597 P-VLPRGADYVSVBELRE--GAVVIKTSLSFYDCVAVTELPSAQCAQCVSSRWGCNW 653  
QY 675 CKYRHVCTH-----DPKTCSPQSGRVK-----LPE-----D 700  
Db 654 CVMQHLCTHAKSCDAGFMVASHQSPVSPDPARGDGSAALEGPPAPLILPESLDYQVD 713  
QY 701 CPOLLRVDR-----ILVPEVKEITLAKKNLPQOSQORGEICLNI 743  
Db 714 TGLMELEATLGASSCPVESVOGSTLMPVHVVEREIRLLGRNLHLFQDGPGBNECMEL 773  
QY 744 QGSEQRVPAL-----RPNSSVQCNTSYSEGBEINNLPELTVVWNGHFNINPAQN 797  
Db 774 EGLEVVVEARVECEPPDPDQCHVTCQHQHLSYEAQ-PELRVGLFLRRAGRLRVDSAEGL 832  
QY 798 KHLKYCGAMRESGLCLKADPDFACGQGP-GQCTLQHCQPAQSOMLELSGAKSKCT 856  
Db 833 HVLYDCSVGHGDCSCOTAMPQYQVCMCEGERPCVTEACG-----EAEAVATQCP 885  
QY 857 NPRITEIIPVTGREGTKVTIRGENGLLEFRDIASHKVAGVCESPVLDVGYIPABQIVC 916  
Db 886 APLIHSVEPLTGVDGTRVINGNLGHQVQVLDGMVTVAGVPCAVDAQEYEVSSSLVC 945  
QY 917 ENGKAPKQAGFVEICV-----AVCRPFMASSQLYFMTLTLSDLKPSRGPMSGGTQ 971  
Db 946 ITG-ASGEVAGATAVEVPGRGVSEHDFAYQDPKVH-----SIFPARGPRAGTR 996  
QY 972 VITITGNLAG--SNVVMFGQPC-LFRRSPSYVCMNTSSDVELEMKVSV-----1021  
Db 997 LTLNGSKLLTGLREDIRVVVVGDPQCHLLPEQOSEQLRCETSPRPTPATLPVAVMFGATER 1056  
QY 1022 QVDRAKIHQDLVFQYVEDPTIVRIEPEWSIVSGNTPIAVMGTHLDLIONPOIRA-----1075  
Db 1057 RLQRCQ-?----FKYLDENITSAGTKFSLGGRICVWGQNLVDVQPIRIVTVVSRM 1110  
QY 1076 -----KHGKKEHINTCEVINATEMTQCAPALALGPDHOSDLT 1112  
Db 1111 LQPSQGLRRRRVVPETACSLGSCSSQCFEPCNVNSSQLITCRTPAL---FGLPEDPW 1167  
QY 1113 ERPEFEGFILDN-VQSLLLILKNTFTYYPNPVPEAFGPSG---ILELPGCTPILKGNL 1168  
Db 1168 VRVE---FTILDNLVDFATLNPFTFSEYADPTLOPLNPEDPTMPFRHKPGSVFSEGENL 1224  
QY 1169 IPPVAGNVKLVTVLVGSKPCTV-TVSDVQLLCESP--NLICRHKVM-----A 1214  
Db 1225 DLAMS-----KEEVAMIGDPCVVKTLLRHLYCEPPVQPLPRHHLREAPDSLPFTV 1280  
QY 1215 RVGMEYSPGMVYIAPDSPLSLP--AIVSTAVAGLLIIFIVAVLIAVYKRSRESDLTK 1272  
Db 1281 QMGNLRFSLGHVQYDQGESPCAFVPAQVGLGVGTSLLALGVIIIVLMYRKSKQALRDYK 1340  
QY 1273 RLQOMQNDLESVALKEKFAELQTDIHELTSLDGAGIPFDYTYTMRVLFPQTEHD 1332  
Db 1341 KVQIQLENLESSVRDCKKEFTDLMTMTDLTSLDLSGSGIPFDLYKYAERIFPFGHRES 1400

QY 1333 PVLRDLEVPYQROERBEKGLKFLAQLINNKFVLLSFRTLESQSFMRDRGNVASLIMT 1392  
Db 1401 PLHRDLGVPSRPTVEQGLQGLSNLNSKFLTKFHTLESQRTFSARDRAYVASLITV 1460  
QY 1393 VLOSKEYATDVILKQLLADLIDKNLESKNPKLLRLRTESVAEKMLTNWTFLLLYKFLKE 1452  
Db 1461 ALHGKLEYFTDILRTLSDLVAQV-AKN-PKMLRRTETVVEKLLTNWMSICLYTFVRD 1518  
QY 1453 CAGEPLPSLCAKQOMEKPIDAITGEARVSLSEDKLIROIDYKTLVLS---CVSPDN 1509  
Db 1519 SVGEPLMYLFRGKHQVKGPDVSVTKAKYTLNDRLLREDVEYRPLTLNALLAVGPGA 1578  
QY 1510 ANSPEVVKILNCDDITQVKEKILDALIFKNVPCSHRPKAADMDLEWROGSGARMLOQED 1569  
Db 1579 GEAGVVPKVLDCDITISQAKEKMLDQYKGVPLTQRPDPRLDVEWRSVAGHLILSD 1638  
QY 1570 ITTKIENDWRLNLTLAHYQVDPGVSVALV---SKOVTAYNANNSTVSRSTASKYENMIR 1626  
Db 1639 VTSEVOGLMRRLTLQHYKVPDGTATVALVPCLTKHVLR-----ENQDY 1681  
QY 1627 YTGSPDLSRSTPMITPDLESQVGMHLVKNHGHGDKQEGDRGS-----KMWSEY 1677  
Db 1682 VPG-----BRTPMLEDVDEGIRPWHLVKPSDEPEPPRPRGRSLRGGERAKAIPY 1735  
QY 1678 LTRLLATKGTLOKFVDDLFETIFSTAHRGSGALPLAIKYMFDLDEQADKHGHDHPVRHT 1737  
Db 1736 LTRLLSMKGTLOKFVDDLFQVILSTSR---PVLAVKYFFLLDEQAQOHGISDQDTHI 1792  
QY 1738 WKNCPLPLRFWNNMKNPQFVDIHKNSITDACLSVVAQTMDSCSTSEHRLKDKSPNK 1797  
Db 1793 WKTNSPLPLRFWINIKNPQFVDQTSNDMDAVLLVIAQTMDACTLADHKLGRDSPINK 1852  
QY 1798 LLYAKOTPSYKNNVRYVSDIGKMPAISDQDMNAYLAEOSRHMNEFTMALSIFSV 1857  
Db 1853 LLYARDIPRYKRMVRYADIRQTVPASDQEMNSVLAELSWMYSGDLGARVALHELYKI 1912  
QY 1858 GYSEERILGLDHDQCGKOKLAYKLEQV 1886  
Db 1913 NKYDQIITALEEDGTAKMQLGYRLQOI 1941

## RESULT 15

O43157 PRELIMINARY; PRT; 2135 AA.  
ID O43157  
AC O43157  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE KIAA0407 protein.  
GN KIAA0407.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN.  
RK MEDLINE=98116655; PubMed=9455477;  
RA Ishikawa K., Nagase T., Nakajima D., Seki N., Ohira M., Miyajima N.,  
RA Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. VIII.  
RT 78 new cDNA clones from brain which code for large proteins in  
RT vitro.";  
RL DNA Res. 4:307-313(1997).  
DR EMBL; AB007867; BAA23703.1; -  
DR InterPro; IPR003006; IG MHC.  
DR InterPro; IPR002909; IPT TIG.  
DR InterPro; IPR000566; LipOcln\_cytFABP.  
DR InterPro; IPR003659; Flexin-like.  
DR InterPro; IPR002165; Flexin\_repeat.  
DR InterPro; IPR001936; RasGAP.  
DR InterPro; IPR001627; Sema.







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 16:04:26 ; Search time 10686 Seconds  
(without alignments)  
5163.665 Million cell updates/sec

Title: US-09-964-956-13  
Perfect score: 1896  
Sequence: 1 MKAMPNWTCLLSHLLMVG.....QKLAYKLEQVITLMSLDSNK 1896

Scoring table:  
OLIGO Xgapop 60.0 , Xgapext 60.0  
Xgapop 60.0 , Ygapext 60.0  
Ygapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US0964956/runat\_24062003\_124828\_7356/app\_query.fasta\_1.2055  
-DB=GenEmbl -QFWT=fastap -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0964956.ecgn\_1.1.7475.\_runat\_24062003\_124828\_7356 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_hgo\_hum.\*  
40: em\_hgo\_mus.\*  
41: em\_hgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	745	39.3	5986	9	AB046770 Homo sapi
2	457	24.1	2020	9	BC028744 Homo sapi
3	457	24.1	3016	9	BC007934 Homo sapi
C 4	396	20.9	179145	9	AC009785 Homo sapi
C 5	396	20.9	188090	9	AC011625 Homo sapi
C 6	385	20.3	3666	9	AK095606 Homo sapi
C 7	383	20.2	179592	2	AC009723 Homo sapi
C 8	252	13.3	1029	9	HSN802037 Homo sapi
C 9	202	10.7	74655	2	AC025595 Homo sapi
10	197	10.4	3556	9	HSN805617 Homo sapi
11	124	6.5	601	6	AX088116 Sequence
12	122	6.4	3560	9	AK091538 Homo sapi
C 13	92	4.9	105810	2	AC131554 Rattus no
C 14	90	4.7	173385	2	AC102171 Mus muscu
C 15	90	4.7	198470	9	AC018643 Human Chr
C 16	90	4.7	222704	2	AC096840 Rattus no
17	81	4.3	776	4	AF176418 Bos tauru
18	78	4.1	152973	2	AC112331 Rattus no
C 19	78	4.1	222704	2	AC096840 Rattus no
C 20	78	4.1	260122	2	AC093959 Rattus no
C 21	71	3.7	173385	2	AC102171 Mus muscu
C 22	71	3.7	185287	9	AC105443 Homo sapi
C 23	66	3.5	199734	2	AC102141 Mus muscu
C 24	65	3.4	239404	2	AC102156 Mus muscu
C 25	63	3.3	5263	9	HSNOVPROT
26	63	3.3	6163	5	XELPLEX
27	63	3.3	6466	10	D86948
28	63	3.3	239404	2	AC102156 Mus muscu
29	56	3.0	75429	2	AC101416 Mus muscu
C 30	56	3.0	152973	2	AC112331 Rattus no
C 31	54	2.8	242	9	HS63MRNAG
32	54	2.8	5982	10	D86950 Mus musculu
33	54	2.8	6039	9	HSSEXGENE
C 34	54	2.8	160656	2	AL592072 Homo sapi
C 35	54	2.8	188873	2	AC073658 Mus muscu
C 36	54	2.8	193961	2	AC025784 Mus muscu
37	54	2.8	198653	2	AC091474 Mus muscu
38	54	2.8	200322	2	AL808145 Mus muscu
C 39	54	2.8	204152	2	AL807376 Mus muscu
C 40	54	2.8	214660	2	AC094668 Rattus no
41	54	2.8	219447	9	HUMFLNG6PD
42	51	2.7	1403	10	BC024509 Mus muscu
43	51	2.7	6730	6	AX305728 Sequence
44	51	2.7	6730	10	D86949 Mus musculu
45	51	2.7	260122	2	AC093959 Rattus no

ALIGNMENTS

RESULT 1

```
AB046770      5986 bp      mRNA      linear      PRI 10-MAY-2002
LOCUS      Homo sapiens mRNA for KIAA1550 protein, partial cds.
DEFINITION      AB046770
ACCESSION      AB046770
VERSION      GI:20521937
KEYWORDS
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
             clone.fh16159s1.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Nagase,T., Kikuno,R., Nakayama,M., Hirose,M., and Ohara,O.
TITLE      Prediction of the coding sequences of unidentified human genes.
             XVIII. The complete sequences of 100 new cDNA clones from brain
             which code for large proteins in vitro
JOURNAL      DNA Res. 7 (4), 273-281 (2000)
MEDLINE      20450683
PUBMED      10997877
REFERENCE      2 (bases 1 to 5986)
AUTHORS      Ohara,O., Nagase,T. and Kikuno,R.
TITLE      Direct Submission
JOURNAL      Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
             Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba
             292-0812, Japan (E-mail:cdna@info.kazusa.or.jp,
             URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
             Fax:81-438-52-3914)
COMMENT      On May 9, 2002 this sequence version replaced gi:10047164.
FEATURES             location/Qualifiers
             1..5986
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="fh16159s1"
             /tissue.type="brain"
             /clone.lib="pBluescriptII SK plus"
             /note="This sequence was obtained by subcloning of the DNA
             fragments derived from two cDNA clones (1 - 1186 was
             derived from p900094 and 1187 - 5986 was derived from
             fh16159).-vector:pBluescriptII SK plus"
             1..5986
             /gene="KIAA1550"
             41..2479
             /gene="KIAA1550"
             /note="Start codon is not identified."
             /codon_start=2
             /product="KIAA1550 protein"
             /protein_id="BAB13376.2"
             /db_xref="GI:20521938"
             /translation="EPRFASSEMKOCVRLTVHPNNTSVSOYNVLVLTYYNVPESL
             VNCTFEDISEMDGLVGNQIOCYSPAKEVPRILITENGDPHVVLQOLKSKETCMFAS
             TSFVFLNCSVHNSCLSCVSPESIRCHWCKTRHCTHDPKTCFQGRVKLPEDCPQLLR
             VDKILYVEVIRKPTLAKNLPPQSGQGRGYECILNIQSGEORVPALRENSSVQCON
             TSYSEGEINNLVELTVVMGHFNIDNPAQNHLYKCGAMRESCGLCLRADPFA
             CNGCGOCTLRQCHPAQESQWLELSGAKSTNPRIETIIPVTPGREGTKVITIRG
             ENGLFEFDIA SHKAVGECPLVDGYTPAQIVCEMGEAKPSOHAGFEICVAVCR
             PFMARSSOLYFMILTLSDLKPSRPMSSGQTQVITITGNLNGSNVVMFGKOPCLF
             HRPSYIVCNTISDQVLEMKVSVQVDRAKIHODLVFQYVEDPTIVRIEPEWSIVSG
             NPTIAVWGTHLDLIDNPQIRAKHGKEHINICEVLNATEMTQAPALALGPDHQSOLT
             ERPEFFGLDNVQSLILNKNTFTYPPNPFAPGSPGILKLPQTPILKGNLIP
             PVAGGNVKNLYTVLVEKPCQTVTSVQQLCESPNLIGRHVMARVGGMEYSFGMYI
             APDPLSLPAIVSIYAVAGLLIIFIVAVLIYKRSRESDLTLKRLQOMQDNLESVA
             LSKGETEPHAGGHCVRVICVCMHICVCVICFIYKQAGAAVGSAGGWRVCVCLC
             ECVCVHVCVCTSVCIIVSTSKOAGQ"
BASE COUNT      1453 a 1611 c 1543 g 1379 t
ORIGIN
Alignment Scores:
Pred. No.:      0
Score:          745.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:    39.29%
DB:             9
Length:         5986
Matches:        745
Conservative:   0
Mismatches:    0
Indels:        0
Gaps:          0
```

Db 1022 TTTCCGCACATCCCTCCCATGTCATCAAGGTTGCTGGGTGAGTGCAGCCCTTTAGTGGAT 1081  
 QY 907 GYTYrileProAlaGluGlnIleValCysGluMetClyGluAlaLysProSerGlnHis 926  
 Db 1082 GGTATCATCCCTGCAGAACAGATCGTGTGAGATGGGGAGGCCAAGCCACCGACAT 1141  
 QY 927 AlaGlyPheValGluIleCysValAlaValCysArgProGluPheMetAlaArgSerSer 946  
 Db 1142 GCAGGCTTCTGGAGATCTCGTGGCTGTGTGCTCGCCTGAATTATGCGCCGCTCTCA 1201  
 QY 947 GlnLeuTyTyPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMet 966  
 Db 1202 CAGCTCTATTACTATGACACATGACTCTCTCAGATCTGAAGCCCGGGGGCCCATG 1261  
 QY 967 SerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnValVal 986  
 Db 1262 TCCGGAGGCCCAAGTGACCATCACAGCACCACTGAATGCCGGAACACAGTGGTG 1321  
 QY 987 ValMetPheClyLysGlnProCysLeuPheHisArgArgSerProSerTyIleValCys 1006  
 Db 1322 GTGATGTTTGGAAAGCAGCCCTGCTCTTCCACAGCGCATCTCCATCTACATTGTCTGC 1381  
 QY 1007 AsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnValAspArgAla 1026  
 Db 1382 AACCCACATCCCTCAGATGAGGTGCTAGAGATGAAGGTGCGGTGAGGTGGACAGGCC 1441  
 QY 1027 LysIleHisGlnAspLeuValPheGlnTyrValGluAspProThrIleValArgIleGlu 1046  
 Db 1442 AAGATCCACAGGACCTGCTTTCAGTATGTGAAGACCCACCATCTGTCGGGATTGAG 1501  
 QY 1047 ProGluTrpSerIleValSerGlyAsnThrProIleAlaValTrpGlyThrHisLeuAsp 1066  
 Db 1502 CCAGAATGAGCATTTCTAGTGAACACACACCCATCGCCGTATGTTGGGACCCACCTGGAC 1561  
 QY 1067 LeuIleGlnAsnProGlnIleArgAlaLysHisGlyClyLysGluHisIleAsnIleCys 1086  
 Db 1562 CTATACAGAACCCCGAGATCCGTCGCAAGATGAGGGAAGGAGACATCAATATCTGT 1621  
 QY 1087 GluValLeuAsnAlaThrGluMetThrCysGlnAlaProAlaLeuAlaLeuGlyProAsp 1106  
 Db 1622 GAGGTTCTGAACGCTACTGAGATGACCTGTGTCAGCGCCGCCCTCGCTGGTCTGTGAC 1681  
 QY 1107 HisGlnSerAspLeuThrGluArgProGluGluPheGlyPheIleLeuAspAsnValGln 1126  
 Db 1682 CACCACTGACACCTGACGAGAGCCCGGAGAGTTTGGCTTCATCTCTGGAACACGTCAG 1741  
 QY 1127 SerLeuLeuIleLeuAsnLysThrAsnPheThrTyrTyrProAsnProValPheGluAla 1146  
 Db 1742 TCCTGCTCTATCTCAACAGACCACTTCACCTACTATCCCAACCCCGGTTTGGGCC 1801  
 QY 1147 PheGlyProSerGlyIleLeuGluLeuLysProGlyThrProIleIleLeuLysGlyLys 1166  
 Db 1802 TTTGGTCCCTCAGGAATCTCTGGAGCTCAAGCCTGGCAGCCCATCATCTCTAAAGGGCAAG 1861  
 QY 1167 AsnLeuIleProValAlaGlyGlyAsnValLysLeuAsnTyrThrValLeuValGly 1186  
 Db 1862 AACCTGATCCCGCTGTGGCTGGGGCAACGTGAAGCTGAACCTGCTGCTGGTGGG 1921  
 QY 1187 GluLysProCysThrValThrValSerAspValGlnLeuLeuCysGluSerProAsnLeu 1206  
 Db 1922 GAGAAGCGTGACCGTGACCGTGTGAGATGTCAGCTGCTCGCGAGTCCCCCAACCTC 1981  
 QY 1207 IleGlyArgHisLysValMetAlaArgValGlyGlyMetGluTyrSerProGlyMetVal 1226  
 Db 1982 ATCGGAGGCACAAAGTAGTGCCCGTGTGCGGTGGCATGAGTACTCCCGGGGATGGTG 2041  
 QY 1227 TyrIleAlaProAspSerProLeuSerLeuProAlaIleValSerIleAlaValAlaGly 1246  
 Db 2042 TACATTGCCCGGACAGCCCGCTCAGCTCCCGGCCATCTGACGATCGCAGTGGCTGGC 2101  
 QY 1247 GluLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrLysArgLysSerArgGlu 1266  
 Db 2102 GGCTCTCTCATATTTTCATCGTGGCGGTGCTCATTCGCTTATAAAGCAAGTCCCGCGAA 2161

QY 1267 SerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAla 1286  
 Db 2162 AGTACCTCAGCTGGAAGCGCTGCAGATGCAGATGGACACCTGGAGTCCCGTGTGGCC 2221  
 QY 1287 LeuGluCysLysGlu 1291  
 Db 2222 CTGAGTGCAAGGAA 2236

## RESULT 2

BC028744  
 LOCUS BC028744 Homo sapiens, similar to plexin protein, clone MGC:34033  
 DEFINITION IMAGE:4828816, mRNA, complete cds.

ACCESSION BC028744  
 VERSION BC028744.1 GI:22382104  
 KEYWORDS MGC.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2020)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission

JOURNAL Submitted (29-APR-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

Tissue Procurement: Miklos Paikowits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 47 Row: b Column: 11  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis.

## FEATURES

Location/Qualifiers  
 1..2020  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:34033 IMAGE:4828816"  
 /tissue\_type="Testis"  
 /clone\_lib="NIH-MGC\_97"  
 /lab\_host="DH10B"  
 /note="vector: pBluescript"  
 383..1951  
 /codon\_start=1  
 /product="similar to plexin protein"  
 /protein\_id="AAH28744.1"  
 /db\_xref="GI:22382105"

## CDS

/translation="MKAMPNWTCLLSHLMVGMSSLTLLTRQAPLSOKORSFVTFR  
 GEPAGFNHLVDEPTGHIYLGAVNRIYKSSDLKVLVTHETGPDENPKCYPRIVQ  
 TCNPLTITNNVKNMLIDYKENRLIACGSLYOGICKLLRLEDFKLGEPYHKRHYL  
 SGVNSCGFVGVISYSLDDKLFATATADGKPEYFPTISSRKLTKNSEADGMPAYVF  
 HDEPVASIKIYSDTFTLIPDFDIYVYFSGNFVYFLLTQPEWSPPGSTTKQGV  
 TSKVLRLCKEDTAFNSVVEYVIGGERSVEYRLQAAYLSKAGAVLGRGLGVHPDDLL  
 LFTVFSQKQKRMKSLDESALCIFKLTINDRIKQYSCYRGEGTDLAWLKVADIP  
 CSSALLTDDNFCGLDMNAPLGVSDMVRGIPVFTEDRDMTSVIAYVYKNISLAFVGT  
 KSKGLKSGFTGPGQGTQEWIGVAGDPGPGANIASQEQMLCVYLQCSSHKHATSQORVQ  
 PLLCCFLNVPNSS"

BASE COUNT 490 a 544 c 450 t  
 ORIGIN

## Alignment Scores:

Pred. No.:	0	Length:	2020
Score:	457.00	Matches:	457
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	24.10%	Indels:	0
DB:	9	Gaps:	0

US-09-964-956-13 (1-1896) x BC028744 (1-2020)

QY	1	MetrysalalaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMet	20
DB	383	ATGAAGCCATGCCTGGAACTGGACCTGCCTCTCTCCACCTCCCTCATGGTGGCAATG	442
QY	21	GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe	40
DB	443	GGCTCCCTCCACTTTGCTACCCGGCAGCCAGCCCGCTGTCCAGAGACGCGTCAATT	502
QY	41	ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr	60
DB	503	GTCACATTCCGAGGAGAGCCGCGGAGGGTTTCAATCACCTGGTGGTGATGAGAGACA	562
QY	61	GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal	80
DB	563	GGACACATTACTTTGGGGGGCGCTCAATCGGATTTTACAAGCTCTCCAGCGACTGAAGTC	622
QY	81	LeuValThrHisGluThrGlyProAspGluAspProLysCysTyrProProArgIle	100
DB	623	TTGTGTGCGCATGAGACAGCGCGGAGGACACCAACCCCAAGTGTATCCACCCGCGCATC	682
QY	101	ValGlnThrCysAsnGluProLeuThrThrAsnAsnValAsnLysMetLeuLeuLe	120
DB	683	GTCCAGACCTGCAATGAGCCCTCACCAACCAACAAATGTCAACAAAGATGCTCTCTATA	742
QY	121	AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu	140
DB	743	GACTACAGGAGACAGGCTGATTCCTGTGGAGCCTGTACCAAGGCATCTGCAAGCTG	802
QY	141	LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu	160
DB	803	CTGAGGCTGGAGGACCTCTTCAAGCTGGGGAGCGCTTATCAAGAAGGAGCAGCTATCTG	862
QY	161	SerGlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAsp	180
DB	863	TCAGGTGTCAAGGAGCGCTCACTTTTGGAGTGATCGCTCTCAGCAACCTGGAT	922
QY	181	AspLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSer	200
DB	923	GACAACTGTTCATGCGCCAGCGCAGTGATGGGAGCCCGAGTATTTCCACCATCTCC	982
QY	201	SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAsp	220
DB	983	AGCGGAAACTGACCAAGAACTCTGAGCGGATGGCATGTTCCGCGTCTCTCCATGAT	1042
QY	221	GluPheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPhe	240
DB	1043	GAGTCTGGGCTCGATGATTAAGATCCTTCGGACCTTCACCATCATCTCCGACTTT	1102
QY	241	AspIleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGln	260
DB	1103	GATATCTACTATGCTATGGTTTATGAGTGGCACTTTGTCTACTTTTGACCCCTCCAA	1162
QY	261	ProGluMetValSerProProGlySerThrThrLysGluGlnValTyrThrSerLysLeu	280
DB	1163	CCTGAGTGGTCTCTCCACGAGCTCCACCAAGAGGAGCGGTGTATACATCCAGGCTC	1222
QY	281	ValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCys	300
DB	1223	GTGAGGCTTTGCAAGGAGACACAGCTTCACTCTATGTAGAGGTGCCCATTTGGCTGT	1282
QY	301	GluArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAla	320

Db	1283	GAGCGCAGTGGGTGGAGTACCGCTGCTGAGCGCTGCTACCTGTCCAAAGGGGGGCC	1342
QY	321	ValLeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSer	340
Db	1343	GTGCTTGGCAGGACCTTGGAGTCCATCCAGATGATGACCTCTTTCACCGCTCTCTCC	1402
QY	341	LysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeu	360
Db	1403	AAGGCCAGAAAGCGGAAATCAATCCCTGATGATGAGTGGCGCTGTGCATCTTCATCTTG	1462
QY	361	LysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThr	380
Db	1463	AAGCAGATAAATGACCGCATTAAGAGCGGCTGCACTTGTACCGGGCGGAGGCGACG	1522
QY	381	LeuAspLeuAlaTrpLeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIle	400
Db	1523	CTGGACCTGGCTGCTCAAGTGAAGGACATCCCTGCGAGTGGCTCTTAACCATTT	1582
QY	401	AspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArg	420
Db	1583	GACGATAACTTCTGTGGCTGGACATGAATGCTCCCTGGAGTGTCCGACATGGTCCGT	1642
QY	421	GlyIleProValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrValTyr	440
Db	1643	GGAATTCCTGCTTCCAGCGAGGACAGGACCGCATGACCTGTCTCATCATATGTCTAC	1702
QY	441	LysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLysLys	457
Db	1703	AAGAACCACCTCTCTGGCTTTGTGGGCACCAAAAGTGGCAAGCTGAAGAAG	1753

RESULT 3  
BC007934/c  
LOCUS  
DEFINITION  
Homo sapiens, DKFZP434A043 protein, clone MGC:14234 IMAGE:4127359,  
mRNA, complete cds.  
ACCESSION  
BC007934  
VERSION  
BC007934.1 GI:14044003  
KEYWORDS  
MGC.  
SOURCE  
Homo sapiens.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1. (bases 1 to 3016)  
Strausberg, R.  
Direct Submission  
Submitted (11-MAY-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nsl.nih.gov](mailto:nisc_mgc@nsl.nih.gov)  
Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, O.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,  
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
Zhang, L.-H. and Green, E.D.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 20 Row: f Column: 2  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF

```
analysis.
FEATURES
  source      Location/Qualifiers
    1..3016
      /organism="Homo sapiens"
      /db_xref="LocusID:25852"
      /db_xref="taxon:9606"
      /clone_lib="NIH_MGC_19"
      /tissue_type="Brain, neuroblastoma"
      /lab_host="DH10B-R"
      /note="Vector: pOTB7"
    CDS
      1490..2563
        /codon_start=1
        /product="DKFZP434A043 protein"
        /protein_id="AAH07934.1"
        /db_xref="GI:14044004"
        /translation="MKNVAVGNKORANLIVLGVAPRLVLYLLOETSSSTELKTECAVV
        LGLAMGTENNYSKSLDCHLIPALLQGLLSPDKLFEACLRCLRTFTSPVTEELY
        TQATVPHLMALLSRKYTOEYICQIFSHCCRGPDHQTILFNHGAONIAHLTSLSY
        KVMQALKCFSVLAFENQVSMYLVNVLFGDELLPQIFVKMLQDRPIEMOLTSACL
        TYMCRAGAITDDNCIVLTPCLVRMCSEKLELVEVEGAETLAVLIEPDVELQRIA
        SITDHLIAMLADYFKYPPSSVSAITDIKRLDHDHLKHAHELQAAFKLYASLGANDEDIR
        KKVSLGGRPPVLTASROGVTSN"
BASE COUNT  742 a  705 c  781 g  788 t
ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3016
Score:          457.00    Matches:      457
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match:      24.10%    Indels: 0
DB:                9      Gaps: 0

US-09-964-956-13 (1-1896) x BC007934 (1-3016)
QY 678 ArgHisValCysThrHisaspProLysThrCysSerPheGlnGluGlyArgValLysLeu 697
Db 1385 CGGCATGTCGACCCATGACCCCAAGACCTGCTCTTCAGGAAGCCGAGTGAAGCTG 1326
QY 698 ProGluAspCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIle 717
Db 1325 CCGGAGGACTGCCCCAGCTGCTGCAGTGGACAAAGATCTCTGTCGCCGTGGAGGTGATC 1266
QY 718 LysProIleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyr 737
Db 1265 AAGCCTATACGCTGAAGGCAAGAACCTCCGCCAGCCAGCTCTGGCGCAGCGTGGCTAC 1206
QY 738 GluCysIleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArgPheAsnSer 757
Db 1205 GAATGCATCTCAACATTCAGGCGAGCGAGCGAGTGCCTGCCCTTCAACAGC 1146
QY 758 SerSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsnLeu 777
Db 1145 TCCAGGCTACAGTGCAGACACCTCTTATCTATGAAGGATGGAGATCAACAACCTG 1086
QY 778 ProValcLeuThrValValTrpAsnGlyHisPheAsnIleAspAsnProAlaGlnAsn 797
Db 1085 CCGCTGGAGTTGACGTGCTGGAATGGCACTTCAACATTTGACAAACCCAGCTCAGAAT 1026
QY 798 LysValHisLeuTyrLysCysGlyAlaMetArgLysSerCysGlyLeuLysAla 817
Db 1025 AAAGTTCCACTCTACAAAGTGTGAGGCATCGTGAGAGCTGCGGGCTGTGCTCAAGGCT 966
QY 818 AspProAspPheAlaCysGlyTrpCysGlnGlyProGlyGlnCysThrLeuArgGlnHis 837
Db 965 GACCCAGACTTCCGATGTGCTGTGCGAGGGCCAGCGAGTGCACCTGCGGCCAGCAC 906
QY 838 CysProAlaGlnGluSerGlnTrpLeuGluLeuSerGlyAlaLysSerLysCysThrAsn 857
Db 905 TGCCTCCGCCAGGAGACCAAGTGGCTGAGCTGCTGGTGCCTCAAAAGCAAGTGCACAAAC 846
QY 858 ProArgIleThrGluIleIleProValThrGlyProArgGluGlyGlyThrLysValThr 877
```

```
|||||
845 CCCGCGATCACAGATAATCCCGTGACAGGCCCGCGGAAGGGGACCAAGGTCAC 786
QY 878 IleArgGlyGluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAla 897
Db 785 ATCCGAGGGGAGAACCTGGGCTGGAATTTCCGACATCGCTCCCATGTCGAAGGTGCT 726
QY 898 GlyValGluCysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGlu 917
Db 725 GCGCTGGAGTGCAGCCCTTTAGTGATGTTACATCCCTGCAGACAGATCGCTGTGTGAG 666
QY 918 MetGlyGluAlaLysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCys 937
Db 665 ATGGGGAGGCAAGCCAGCCAGCATCGAGCTTCTGCGAGATCTGCGTGTGTGTGT 606
QY 938 ArgProGluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSer 957
Db 605 CGGCTGAATTCATGCGCCGCTCTCACAGCTCTATTACTTACATGACATGACTCTCTCA 546
QY 958 AspLeuLysProSerArgGlyProMetSerGlyGlyThrGlnValThrIleThrGlyThr 977
Db 545 GATCTGAAGCCAGCGGGGCCCATGTCCGGAGGGACCAAGTGACCATCATCAGGCACC 486
QY 978 AsnLeuAsnAlaGlySerAsnValValMetPheGlyLysGlnProCysLeuPheHis 997
Db 485 AACCTGAATGCCGAAGCAACGTTGGTGTGATGTTTGGAAAGCAGCCCTGCTCTTCCAC 426
QY 998 ArgArgSerProSerTyrIleValCysAsnThrThrSerSerAspGluValLeuGluMet 1017
Db 425 AGCGCATCTCCATCTACATTTCTGCAACACCATCTCTCAGATGAGGTGCTAGAGATG 366
QY 1018 LysValSerValGlnValAspArgAlaLysIleHisGlnAspLeuValPheGlnTyrVal 1037
Db 365 AAGGTGTCGGTGCAGTGGACAGGCGCAAGATCCACAGGACCTGGTCTTTTCAGTATGTG 306
QY 1038 GluAspProThrIleValArgIleGluProGluTrpSerIleValSerGlyAsnThrPro 1057
Db 305 GAAGACCCCATCTGTCGGATTGAGCCAGAAATGAGCATTTGTCAGTGGAAACACACC 246
QY 1058 IleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHis 1077
Db 245 ATCCCGCTGTGGGGACCCACCTGGACCTCATACAGAACCCACCATCGCTGCCAAGCAT 186
QY 1078 GlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGln 1097
Db 185 GGAGGAAGGAGCACATCAATATCTGTGAGGTCTTGAACGCTACTGAGATGACCTGTCTAG 126
QY 1098 AlaProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGlu 1117
Db 125 GCGCCCGCCCTCGCTCTGGGTCTGACCACCATGTCAGACCTGACCGAGAGGCCCGAGGAG 66
QY 1118 PheGlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThr 1134
Db 65 TTTGGCTTCATCTCGGACAAACGTCAGTCCCTGCTCATCTCTCAACAAGACC 15

RESULT 4
AC009785/c 179145 bp DNA linear PRI 26-MAY-2002
LOCUS Homo sapiens chromosome 7 clone RP11-198E23, complete sequence.
DEFINITION AC009785
ACCESSION AC009785
VERSION AC009785.8 GI:21217401
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 179145)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 179145)
```

**AUTHORS** Bubba, K.L., Desmarais, C.L., Ramsey, S.A. and Hubley, R.M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (01-SEP-1999) Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
**REFERENCE** 3 (bases 1 to 179145)  
**AUTHORS** Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Raymond, C., Clendenning, J., Ivey, R.G. and Haugen, E.D.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (20-JUN-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
**REFERENCE** 4 (bases 1 to 179145)  
**AUTHORS** Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (06-APR-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
**REFERENCE** 5 (bases 1 to 179145)  
**AUTHORS** Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (26-MAY-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
**COMMENT** On May 26, 2002 this sequence version replaced gi:20066281.  
-----  
**Center:** University of Washington Genome Center  
**Center Code:** UWGC  
**Web site:** <http://www.genome.washington.edu>  
**Contact:** uwgctgs@u.washington.edu  
-----  
**Project Information**  
**Center project name:** chr-7  
**Center clone name:** RP11-198E23 (djs380)  
-----  
**Summary Statistics**  
**Sequencing vector:** plasmid; X52328; 100% of reads  
**Chemistry:** Dye-terminator Big Dye; 100% of reads  
**Assembly program:** Phrap; version 0.990319  
**Consensus quality:** 178850 bases at least Q40  
**Consensus quality:** 179053 bases at least Q30  
**Consensus quality:** 179141 bases at least Q20  
**Insert size:** 179145; sum-of-contigs  
**Quality coverage:** 11.8x in Q20 bases; sum-of-contigs  
-----  
**Overlapping Sequences:**  
5': RP11-341C17 (UWGC:djs301) AC011625 73322-bp overlap  
3': RP11-335B7 (UWGC:djs542) AC026239 38206-bp overlap  
-----  
**Sequence Quality Assessment:**  
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.  
-----  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.  
-----  
**Sequence Validation:**  
This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies

between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

SeqDerMap	ECORI	FngPrnt	SeqDerMap	HindIII	FngPrnt	SeqDerMap	BglII
-----	-----	-----	-----	-----	-----	-----	-----
8696	8850	<800	399	<800	2304	2280	
6	<800	6347	6382	6347	2067	2069	
1907	1895	<800	512	<800	5518	5637	
1854	1895	<800	449	<800	6918	7002	
2849	2922	2276	2322	2276	2029	2069	
3722	3806	5995	5951	5995	2378	2425	
658	<800	7291	7224	7291	2089	2069	
26398	26593	1033	1005	1033	550	<800	
537	<800	1033	1020	1033	4191	4187	
1542	1483	3448	3445	3448	3839	3773	
1103	1096	2126	2131	2126	1647	1646	
1928	2045	2276	2237	2276	629	<800	
14029	13827	6652	6625	6652	1932	1918	
6707	6816	6347	6349	6347	18266	18041	
1884	1895	6652	6669	6652	3099	3165	
6048	6029	1916	1943	1916	2900	2928	
2150	2209	6347	6376	6347	111	<800	
13944	13827	5641	5398	5641	2425	2425	
4275	4287	<800	599	<800	3767	3773	
2057	2209	4127	4055	4127	1506	1485	
2293	2209	<800	299	<800	3639	3773	
6448	6393	20079	20065	20079	5635	5637	
9916	9943	6903	6964	6903	6894	7002	
7673	7645	5641	5766	5641	11819	11807	
115	<800	4127	4281	4127	837	821	
1490	1483	7901	7844	7901	2058	2069	
1071	1096	1033	1032	1033	3280	3312	
18341	18612	1546	1591	1546	3780	3773	
7618	7645	954	955	954	623	<800	
1733	1700	<800	330	<800	202	<800	
1478	1483	1916	1927	1916	2950	2928	
7616	7645	6347	6261	6347	10775	10621	
561	<800	<800	445	<800	12114	11807	

-----	19080	18612	-----	3869	3877	-----	9445	9461
-----	114	<800	-----	4148	4127	-----	1694	1646
-----	7	<800	-----	3828	3877	-----	3745	3773
-----			-----	4001	4127	-----	674	<800
-----			-----	1561	1546	-----	9893	9817
-----			-----	3188	3192	-----	3225	3165
-----			-----	4960	4943	-----	998	1011
-----			-----	9109	8931	-----	1318	1283
-----			-----	929	954	-----	94	<800
-----			-----	7844	7901	-----	796	821
-----			-----	535	<800	-----	111	<800
-----			-----	5397	5408	-----	1029	1011
-----			-----	1905	1916	-----	1491	1485
-----			-----	113	<800	-----	20564	20918
-----			-----	5596	5641	-----		
-----			-----	750	735	-----		
-----			-----	1264	1248	-----		
-----			-----			-----		
FEATURES	Location/Qualifiers							
source	1..179145							
	/organism="Homo sapiens"							
Alignment Scores:								
Pred. No.:	0	Length:	179145					
Score:	396.00	Matches:	396					
Percent Similarity:	100.00%	Conservative:	0					
Best Local Similarity:	100.00%	Mismatches:	0					
Query Match:	20.89%	Indels:	0					
DB:	9	Gaps:	0					
US-09-964-956-13 (1-1896) x AC009785 (1-179145)								
Qy	1	MetLysAlaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMet	20					
Db	61796	ATGAAGCCATGGCCTGGAACCTGGACCTGCTTCTCTCCACCTCCCTCATGGTGGCATG	61737					
Qy	21	GlySerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe	40					
Db	61736	GGTCTCTCCACTTGTCTCAACCCGCGACCCGCTGCTCCCAAGAGCGGTCATTT	61677					
Qy	41	ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr	60					
Db	61676	GTCACATTCCGAGAGAGCCGCGAGGGTTCAATCACCTGGTGGTGATGAGAGACAC	61617					
Qy	61	GlyHisThrLeuGlyAlaValAsnArgIleTrpLysLeuSerSerAspLeuLysVal	80					
Db	61616	GGACACATTTACTTGGGGCGGTCAATCGGATTTTACAAAGCTCTCCAGCGACCTCAAGGTC	61557					
Qy	81	LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle	100					
Db	61556	TTGGTGACGATGAGACGGCGCGAGAGCAACCCCAAGTTTACCCACCCCGCATC	61497					
Qy	101	ValGlnThrCysAsnGluProLeuThrThrAsnAsnValAsnLysMetLeuLeuLe	120					
Db	61496	GTCAGACCTGCAATGAGCCCTGACCACCACCAACATGTCAACAAGATGCTCTCAT	61437					

Qy	121	AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu	140
Db	61436	GACTACAGGAGACAGGCTGATTGCTCTGGAGCCTGTACCAAGGCATCTGCAAGCTG	61377
Qy	141	LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu	160
Db	61376	CTGAGGCTGGAGGACCTCTCAAGCTGGGGAGCCTTATCATAGAAGAGGACCATATCTG	61317
Qy	161	SerGlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAsp	180
Db	61316	TCAGGTGTCAACGAGAGCGGCTCAGTCTTTGGAGTGATCGTCTCTACAGCAACCTGGAT	61257
Qy	181	AspLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSer	200
Db	61256	GACAAGCTGTTCATTCGCCAGCGCAGTGGATGGGAAGCCGAGTATTTTCCACCATCTCC	61197
Qy	201	SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAsp	220
Db	61196	AGCCGAACCTGACCAAGAACTCTGAGCGGATGGCATGTTCCGGTACCTCTTCCATGAT	61137
Qy	221	GluPheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPhe	240
Db	61136	GAGTTCGTGGCTCGATGATTAAAGATCCCTTCGGACACCTTCACCATCATCCTGACTTT	61077
Qy	241	AspIleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGln	260
Db	61076	GATATCTACTATGCTATGGTTTATAGCAGTGGCAACTTTTGTCTACTTTTGTACCCCTCAA	61017
Qy	261	ProGluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeu	280
Db	61016	CCTGAGATGGTGTCTCCACAGGCTCCACCAAGGACAGGTGTATACATCAAGCTC	60957
Qy	281	ValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCys	300
Db	60956	GTGAGGCTTTCGAAGGAGGACACAGCTTCACTCTATGTAGAGGTGCCCATTTCTCTCC	60897
Qy	301	GluArgSerGlyValGluTyrArgLeuLeuGlnAlaTyrLeuSerLysAlaGlyAla	320
Db	60896	GAGCGACGTGGGTGGAGTACCGCTGCTGCGAGGTGCTTGTCTGTCCAAAGCGGGGCC	60837
Qy	321	ValLeuGlyArgThrLeuGlyValHisProAspAspAspLeuLeuPheThrValPheSer	340
Db	60836	GTGCTTGGCAGGACCCCTTGGAGTCCATCCAGATGATGACCTCTTCTTCCCGCTTCTCTCC	60777
Qy	341	LysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeu	360
Db	60776	AAGGCCAGAGAGCGGAAATGAATCCCTGGATGATGCGCCCTGTGCATCTTCACTCTG	60717
Qy	361	LysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThr	380
Db	60716	AAGCAGATAAATGACCGCATTAAGGAGCGGCTGCAGTCTTGTACCGGGCGAGGCGACG	60657
Qy	381	LeuAspLeuAlaTrpLeuLysValLysAspIleProCysSerSerAla	396
Db	60656	CTGACCTGGCTGCTCAAGGTGAAGACATCCCTCGACGAGTGCG	60609
RESULT 5			
AC011625/c			
LOCUS	AC011625	188090 bp	DNA linear PRI 08-DEC-1999
DEFINITION	Homo sapiens clone RP11-341C17 from 7q31, complete sequence.		
ACCESSION	AC011625		
VERSION	AC011625.2	GI:6539285	
KEYWORDS	HTG.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Kaul,R.K., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D., Green,P.		
JOURNAL	and Olson,M.V.		
REFERENCE	Large-scale MCD Mapping and Sequencing of Human Chromosome 7		
	Unpublished		
	2 (bases 1 to 188090)		

AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Bubb, K.L., Desmarais, C.L., Ramsey, S.A. and Hubley, R.M.  
Direct Submission  
Submitted (08-OCT-1999) Human Genome Center, University of  
Washington, Box 352145, Seattle, WA 98195, USA  
3 (bases 1 to 188090)  
Kaul, R.K. and Desmarais, C.L.  
Direct Submission  
Submitted (08-DEC-1999) Human Genome Center, University of  
Washington, Box 352145, Seattle, WA 98195, USA  
On Dec 8, 1999 this sequence version replaced gi:6016750.

----- Genome Center:  
University of Washington Genome Center  
Center code: UWGC  
Web site: <http://genome.washington.edu>  
Contact: [uwgchelp@u.washington.edu](mailto:uwgchelp@u.washington.edu)  
----- Project Information  
Center project name: chr-7  
Center clone name: djs301 (RP11-341C17)  
----- Summary Statistics  
Sequencing vector: M13; 100% of reads  
Chemistry: Dye-primer Bodipy; 93% of reads Chemistry:  
Dye-terminator Big Dye; 7% of reads  
Assembly program: Phrap; version 0.990319  
Insert size: 188 094; sum-of-contigs  
Quality coverage: 7.78X in Q20 bases; sum-of-contigs

----- Overlapping Sequences:  
5': UWGC:djs302  
3': UWGC:djs380  
-----  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
GenBank flat file format but are available as part  
of this entry's ASN.1 file.

Double stranded (DS) coverage: 70.4%  
DS or two chemistry coverage: 100.0%  
Single stranded regions: 0

----- Sequence Validation:  
This sequence has been validated by Multiple Complete Digest  
fingerprinting. Comparison of the experimentally derived digest  
fragments with sequence-predicted fragments is given below.  
The electronically-digested sequence consists of both insert and  
vector, in order to accurately represent the entire circular BAC.  
Small fragments below a variable cutoff (approximately 400-600 bp)  
are not resolved in the fingerprint and hence do not appear  
in the table. There are no significant remaining discrepancies  
between the experimental and predicted values. Uniquely ordered  
fragments are separated by dashed lines.

BgIII		EcoRI		NsiI	
FP	Seq	FP	Seq	FP	Seq
2095.00	2061.00	8947.00	8687.00	729.00	732.00
12518.00	12347.00	844.00	812.00	1463.00	1480.00
5082.00	5007.00	4657.00	4593.00	5528.00	5438.00
2694.00	2672.00	3637.00	3613.00	3383.00	3371.00
5175.00	5083.00	1152.00	1147.00	1240.00	1220.00
1094.00	1093.00	11517.00	11411.00	4875.00	4844.00
2231.00	2189.00	16729.00	16537.00	4194.00	4094.00

16006.00	15888.00	13550.00	13383.00	2402.00	2353.00
6982.00	6668.00	5001.00	4900.00	1463.00	1439.00
8822.00	8629.00	6353.00	6107.00	2072.00	2056.00
11945.00	11811.00	3101.00	3027.00	5301.00	5143.00
10740.00	10587.00	3850.00	3719.00	1463.00	1408.00
2095.00	2062.00	1925.00	1935.00	729.00	706.00
7981.00	7820.00	6169.00	6045.00	10419.00	10239.00
2095.00	2101.00	6353.00	6206.00	4194.00	4107.00
1178.00	1141.00	1743.00	1714.00	1128.00	1118.00
3623.00	3525.00	2067.00	2052.00	622.00	630.00
621.00	581.00	559.00	569.00	1045.00	1052.00
6428.00	6341.00	544.00	510.00	7407.00	7153.00
6148.00	6023.00	2364.00	2354.00	3498.00	3389.00
769.00	736.00	660.00	633.00	4506.00	4428.00
4722.00	4640.00	1013.00	972.00	2402.00	2339.00
1178.00	1180.00	11962.00	11788.00	13094.00	12940.00
7083.00	6917.00	5201.00	5141.00	4378.00	4306.00
2095.00	2022.00	5201.00	5091.00	2025.00	1988.00
2429.00	2372.00	1925.00	1902.00	4037.00	3911.00
2095.00	2083.00	1925.00	1853.00	2402.00	2326.00
546.00	544.00	2967.00	2843.00	2072.00	2069.00
4240.00	4182.00	3850.00	3715.00	1609.00	1583.00
3913.00	3829.00	560.00	652.00	2922.00	2845.00
1656.00	1641.00	29666.00	26384.00	5301.00	5252.00
621.00	623.00	544.00	531.00	1747.00	1719.00
1940.00	1927.00	1551.00	1536.00	729.00	697.00
18696.00	18261.00	1152.00	1097.00	2646.00	2580.00
3153.00	3085.00	1925.00	1922.00	2402.00	2335.00
2966.00	2898.00	14113.00	14024.00	10419.00	10108.00
2429.00	2419.00	6876.00	6704.00	7407.00	7317.00
3913.00	3763.00	1925.00	1869.00	8785.00	8640.00
1524.00	1500.00	6169.00	6048.00	7052.00	6870.00
3776.00	3636.00	2207.00	2144.00	3731.00	3659.00
5704.00	5638.00			4194.00	4085.00
6982.00	6868.00			622.00	581.00
				3881.00	3751.00
				1240.00	1221.00

-----  
 1319.00 1284.00  
 -----  
 5528.00 5416.00  
 -----  
 7052.00 6844.00  
 -----  
 3881.00 3888.00  
 -----  
 3498.00 3426.00  
 -----  
 14516.00 14342.00  
 -----

## FEATURES

## source

Location/Qualifiers  
 1. 188090  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="7"  
 /map="7q31"  
 /clone="RP11-341c17 or UWGC:djs301"  
 /cell\_line="Male Blood"  
 /clone\_lib="RPC-11 Human Male BAC Library"  
 444.750  
 /rpt\_family="Alu"  
 891.963  
 /standard\_name="SWS2952"  
 /notes="Genbank Accession: G13043"  
 complement(2131..2409)  
 /rpt\_family="Alu"  
 complement(8590..8831)  
 /rpt\_family="Alu"  
 10548..10834  
 /rpt\_family="Alu"  
 11525..11795  
 /rpt\_family="Alu"  
 18103..18403  
 /rpt\_family="Alu"  
 complement(27089..27369)  
 /rpt\_family="Alu"

## Alignment Scores:

Pred. No.: 0 Length: 188090  
 Score: 396.00 Matches: 396  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 20.89% Indels: 0  
 DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x AC011625 (1-188090)

Qy 1 MetLysAlaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuMetValGlyMet 20  
 |||||  
 Db 176564 ATGAAGCCATGCCCTGGAACTGGACCTGCTCTCTCCACCTCCTCATGGTGGGCATG 176505  
 |||||  
 Qy 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
 |||||  
 Db 176504 GGCTCTCCACTTGTCTACCCGGCGAGCCAGCCGCTGCCAGAGACGCGTCAATT 176445  
 |||||  
 Qy 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
 |||||  
 Db 176444 GTCACATTCGGAGACAGCCGCCGAGGGTTTCATCACCTGGTGGGATGAGAGACA 176385  
 |||||  
 Qy 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
 |||||  
 Db 176384 GGACACATTTACTTGGGGCGCGTCAATCGGATTTACAAGCTCTCCAGCGACCTGAAGTGC 176325  
 |||||  
 Qy 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProArgIle 100  
 |||||  
 Db 176324 TTGGTGACGATGAGACAGCGCGGAGGACCAACCCCAAGTGTATACCCACCCCGCATC 176265  
 |||||  
 Qy 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuLe 120  
 |||||  
 Db 176264 GTCCAGACCTGCANTGAGCCCTTGACCAACCAATGTCAACAAGATGCTCCTCAT 176205

Qy 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
 |||||  
 Db 176204 GACTACAGGAGAACAGAGCTGATTCCTGCTGGAGCCCTGTACCAAGGCATCTGCAAGCTG 176145  
 |||||  
 Qy 141 LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu 160  
 |||||  
 Db 176144 CTGAGGCTGGAGAGCTCTTCAAGCTGGGGAGCCCTTATCATTAAGAAGGAGCACTATCTG 176085  
 |||||  
 Qy 161 SerGlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAsp 180  
 |||||  
 Db 176084 TCAGGTGTCAACGAGAGCGGCTCGGTCTTTGGAGTGATCGTCTCTTACAGCAACCTGGAT 176025  
 |||||  
 Qy 181 AspLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSer 200  
 |||||  
 Db 176024 GACAAGCTGTTTCATTTGCCACGCACTGTGATGGGAAGCCGAGCTATTTCCCAACCATCTCC 175965  
 |||||  
 Qy 201 SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAsp 220  
 |||||  
 Db 175964 AGCCGGAACCTGACCAAGAACTCTGAGGCGATGGCATGTTCCGCTACGCTTCTCCATGAT 175905  
 |||||  
 Qy 221 GluPheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProaspPhe 240  
 |||||  
 Db 175904 GAGTTCTGGGCTCGATGATTAAGATCCCTTCGGACACCTTCACCATCATCCTGACTTT 175845  
 |||||  
 Qy 241 AspileTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGln 260  
 |||||  
 Db 175844 GATATCTACTATGCTATGTTTATAGCAGTGCCAACTTGTCTACTTTTGGACCTCCAA 175785  
 |||||  
 Qy 261 ProGluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeu 280  
 |||||  
 Db 175784 CCTGAGATGGTGTCTCCACAGGCTCCACCAAGAGAGCAGGTGTATACATCCAAAGCTC 175725  
 |||||  
 Qy 281 ValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCys 300  
 |||||  
 Db 175724 GTGAGGCTTTGCCAAGGAGGACACACCTTCAACTCTATGTAGAGGTGCCAATGGCTGT 175665  
 |||||  
 Qy 301 GluArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAla 320  
 |||||  
 Db 175664 GAGCGACGTGGGTGGAGTACCGCTGTGTCAGGCTGCTACCTGTCCAAAGCGGGGCC 175605  
 |||||  
 Qy 321 ValLeuGlyArgThrLeuGlyValHisProaspAspLeuLeuPheThrValPheSer 340  
 |||||  
 Db 175604 GTGCTTGGCAGGACCTTTGGAGTCCATCCAGATGATGACCTCTCTTACCGCTCTCTCC 175545  
 |||||  
 Qy 341 LysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeu 360  
 |||||  
 Db 175544 AAGGCCCAAGCGGAAATCCCTGGATGATGAGTGGCCCTGTGCATCTTCATCTTG 175485  
 |||||  
 Qy 361 LysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThr 380  
 |||||  
 Db 175484 AAGCAGATAAATGACCGCATTAAGGAGCGGCTGCGAGTCTTGTACCGGGCGAGGCGACG 175425  
 |||||  
 Qy 381 LeuaspLeuAlaTrpLeuLysValLysAspIleProCysSerSerAla 396  
 |||||  
 Db 175424 CTGGACCTGGCTGGCTTAAAGGTGAAGGACATPCCCTGCAGCAGTGGC 175377  
 |||||

## RESULT 6

## AK095606

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

AK095606 3666 bp mRNA linear PRI 15-JUL-2002  
 Homo sapiens cDNA FLJ38287 fis, clone FCBBF3008362, moderately  
 similar to PLEXIN 4 PRECURSOR.

AK095606  
 AK095606.1 GI:21754899  
 oligo capping: fis (full insert sequence)  
 Homo sapiens fetal brain cDNA to mRNA, clone\_lib:FCBBF3  
 clone:FCBBF3008362.

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1  
 Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T.,  
 Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R.,

Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Terashima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 3666)  
Isogai, T. and Yamamoto, J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES  
Location/Qualifiers  
1..3666  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="FCBBF3008362"  
/tissue="type="brain"  
/clone\_lib="FCBBF3"  
/dev\_stage="fetal"  
/note="cloning vector: pME18SFL3"  
125..1159  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAC04587.1"  
/db\_xref="GI:21754900"  
/translation="MDLEWRQSGARMILQEDITTKIENDKRLNTLAHVOPDGSV  
VALVSKQVATYNAVNNSTVRSASKYENIRYTGSPSLRTPMIFPDLESQVNMW  
HLVKNHEDQKDRGSKMSEIYTLRLATKGLQKFDVDFETFEIESTAHRGSLP  
LAIKYMFDELQADKHGHIDPHVHTWKNCLPLRFVWMMIKNPFQFDIHKNSITD  
ACLVSVAOTDMSCTSEHRLCKDPSNKLKYAKDIPSYKNWERYSDIGKWPASID  
ODMNAYLAEQSRMHNNFNTNSALSEIFSYGVKYSIEILGPLDHDHQQCKQLAYKLE  
QVITLMSLDS"  
BASE COUNT 944 a 946 c 921 g 855 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0 Length: 3666  
Score: 395.00 Matches: 385  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.31% Indels: 0  
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x AK095606 (1-3666)

QY 1510 AlaAsnSerProGluValProValLysIleLeuAsnCysAspThrIleThrGlnValLys 1529  
Db 2 GCCAACAGCCCGGAGTCCAGTCAAGATCCTCACTGTGACACCATCACTCAGGTCAAG 61

QY 1530 GluLysIleLeuAspAlaIlePheLysAsnValProCysSerHisArgProLysAlaAla 1549  
Db 62 GAGAGATTCTGGATGCCATCTTCAAGAATGTGCCTTGTCTCCACCCGCCCAAGCTGCA 121

QY 1550 AspMetAspLeuGluTrpArgGlnGlySerGlyAlaArgMetIleLeuGlnAspGluAsp 1569  
Db 122 GATATGGATCTGGAGTGGCGACAAGGAGTGGGGCAAGGATGATCTTGGAGGATGAAGAC 181

QY 1570 IleThrThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHisTyrGlnVal 1589

182 ATCCACCACCAAGATTGAGAATGATTGGAAGCGACTGAACACACTGGCCCCACTCAGAGTG 241  
1590 ProAspGlySerValValAlaLeuValSerLysGlnValThrAlaThrAsnAlaValAsn 1609  
242 CCAGATGGTTCGGTGGTGCATTAGTGTCCAAGCAGGTGACACCCCTATAACGAGTGAAC 301  
1610 AsnSerThrValSerArgThrSerAlaSerLysTyrGluAsnMetIleArgTyrThrGly 1629  
302 AACTCCACCGTCTCCAGGACCTCAGCAAGTAATATGAAACATGATCCGGTACACGGGC 361  
1630 SerProAspSerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyVal 1649  
362 AGCCCCGACAGCGCTCCGCTCAGCGACACCTATGATCACTCTCCCTGAGAGTGGAGTC 421  
1650 LysMetTrpHisLeuValLysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGly 1669  
422 AGATGTGGCACCCTAGTGAAGAACACGACGACGAGACCCAGAGAGGGGGGCGGG 481  
1670 SerLysMetValSerGluIleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeuGln 1689  
482 ACCAAGATGGTGTCTCAAAATCTACCTGACCCGACTCTCTGGCCACTAAGGGCACACTCG 541  
1690 LysPheValAspAspLeuPheGluThrIlePheSerThrAlaHisArgGlySerAlaLeu 1709  
542 AAGTTTGTGGATGACCTCTTTGAGACCATCTTCAGCACGCGCACCCCTGCTGCCCTG 601  
1710 ProLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIle 1729  
602 CCCCTGGCCATCAAGTACATGTTTGTACTCTCGATGAGCAGGCTGATAAACATGGCAT 661  
1730 HisAspProHisValArgHisThrTrpLysSerAsnCysLeuProLeuArgPheTrpVal 1749  
662 CATGACCCGACGCTCCGCCATACCTGGAAGAGCAATTCCTGCCCTTGAGGTTTGGGTC 721  
1750 AsnMetIleLysAsnProGlnPheValPheAspIleHisLysAsnSerIleThrAspAla 1769  
722 AACATGATCAAGAACCCGCGAGTTTGTGTTGACATCCATAAGAACACATCAGACAGCC 781  
1770 CysLeuSerValValAlaGlnThrPheMetAspSerCysSerThrSerGluHisArgLeu 1789  
782 TGCCTCTCTGTGGTGCCTCAGACCTTCATGAGCTCTTGTCTCCAGTCAGACACCGGTC 841  
1790 GlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysAspIleProSerTyrLysAsn 1809  
842 GCGAAGGACTCGCCCTCCCAAGCTCTGTATGCCAAGGACATCCCACGCTACCAAGAT 901  
1810 TrpValGluArgTyrTrpSerAspIleGlyLysMetProAlaIleSerAspGlnAspMet 1829  
902 TGGGTGGAGAGGTATTACTCAGACATAGGAAGATGCCAGCCATCAGGACCAAGACATG 961  
1830 AsnAlaTyrLeuAlaGluGlnSerArgMetHisMetAsnGluPheAsnThrMetSerAla 1849  
962 AACGCATACCTGCTGAGCAGTCCCGGATGCACATGAATGAGTTCAACACCATGAGTGA 1021  
1850 LeuSerGluIlePheSerTyrValGlyLysTyrSerGluGluIleLeuGlyProLeuAsp 1869  
1022 CTCTCAGAGATCTTCTCTATGTGGCAAAATACAGCGAGGAGATCTCTGGACCTTGGAC 1081  
1870 HisAspAspGlnCysGlyLysGlnLysLeuAlaTyrLysLeuGluGlnValIleThrLeu 1889  
1082 CACGACGACGAGTGTGGGAAGACGAGAACTGGCCTACAACTAGAACAGTCAACCTC 1141  
1890 MetSerLeuAspSer 1894  
1142 ATGAGCTTAGACAGC 1156

RESULT 7  
AC009723/c 179592 bp DNA linear HTG 08-NOV-2000  
LOCUS Homo sapiens clone RP11-44F19, WORKING DRAFT SEQUENCE, 25 unordered  
DEFINITION pieces.  
AC009723  
AC009723

# VERSION KEYWORDS SOURCE ORGANISM

AC009723.4 GI:965551  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
Homo sapiens

## REFERENCE

1. (bases 1 to 179592)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Depayre, E., Devon, K., Dewar, K., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Toruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

## AUTHORS

## JOURNAL

## REFERENCE

## AUTHORS

2. (bases 1 to 179592)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Depayre, E., Devon, K., Dewar, K., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Toruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

misc\_feature 24348..26991  
/note="assembly\_fragment"  
misc\_feature 27092..31443  
/note="assembly\_fragment"  
misc\_feature 31544..37218  
/note="assembly\_fragment"  
misc\_feature 37319..40338  
/note="assembly\_fragment"  
misc\_feature 40439..45330  
/note="assembly\_fragment"  
misc\_feature 45431..53182  
/note="assembly\_fragment"  
misc\_feature 53283..61147  
/note="assembly\_fragment"  
misc\_feature 61248..73146  
/note="assembly\_fragment"  
misc\_feature 73247..82900  
/note="assembly\_fragment"  
misc\_feature 83001..139568  
/note="assembly\_fragment"  
misc\_feature 139669..179592  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right"  
BASE COUNT 47805 a 41712 c 41281 g 46388 t 2406 others  
ORIGIN

Alignment Scores:  
Pred. No.: 0 Length: 179592  
Score: 383.00 Matches: 396  
Percent Similarity: 99.75% Conservative: 0  
Best Local Similarity: 99.75% Mismatches: 0  
Query Match: 20.20% Indels: 1  
DB: 2 Gaps: 0

US-09-964-956-13 (1-1896) x AC009723 (1-179592)

Qy 1 MetLysAlaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMet 20  
Db 97654 ATGAAAGCCATGCTCGACCTGGACCTGCTCTCTCCACCTCCTCATGTGGGCATG 97595  
Qy 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
Db 97594 GCCTCTCCACTTTCCTCACCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 97535  
Qy 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
Db 97534 GTCACATTCGAG 97475  
Qy 61 GlyHisLleTyrLeuGlyAlaValAsnArgLleTyrLysLeuSerSerAspLeuVal 80  
Db 97474 GGACACATTTACTTGGGGCCGCTCAATCGGATTTTACAGCTCTCCAGCGACCTGAAGTC 97415  
Qy 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgLle 100  
Db 97414 TTGGTGACCATGAGACAGGGCCGACGAGGAGACACCCCAAGTGTACCAACCCGCATC 97355  
Qy 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeu 120  
Db 97354 GTCCAGACCTGCAATGAGCCCTGACCAACCAACCAATGTCAACAAGATGCTCTCAT 97295  
Qy 121 AspTyrLysGluAsnArgLeuLeuAlaCysGlySerLeuTyrGlnGlyLleCysLysLeu 140  
Db 97294 GACTCAAGGAG 97235  
Qy 141 LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu 160  
Db 97234 CTGAGGCTGGAGGACCTCTTCAGCTGGGGAGCCCTATCATAGAAGGAGGACCTATC 97175  
Qy 161 SerGlyValAsnGluSerGlySerValPheGlyValLleValSerTyrSerAsnLeuAsp 180  
Db 97174 TCAGGTGTCAACGAG 97115

Qy 181 AspLysLeuPheLleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSer 200  
Db 97114 GACAAGCTGTTGATTCACCGCAGTGATGGAGAGCCGAGTATTTCCACCATCTCC 97055  
Qy 201 SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAsp 220  
Db 97054 AGCCGGAACACTGACCAAGAACTCTGAGCGGATGGCATGTTGCGGTACGCTTCCATGAT 96995  
Qy 221 GluPheValAlaSerMetIleLysLysLysProSer-AspThrPheThrIleIleProAspPh 240  
Db 96994 GAGTTGCGGGCTCGATGATTAAAGATCCCTTCGGGACACTTCACCATCATCTGACTT 96935  
Qy 240 eAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuG 260  
Db 96934 TGATATCTACTATGCTATGTTTATAGCAGTGCAACTTGTCTACTTTTGGACCTCCA 96875  
Qy 260 nProGluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLe 280  
Db 96874 ACCTGAGATGGTGTCTCCACAGGCTCCACCAAGGAGCAGGTGTATACATCCAAGCT 96815  
Qy 280 uValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProLleGlyCy 300  
Db 96814 CTTGAGGCTTTGCAAGGAGGACACACGCTTCACTCTATGTAGAGTGCCCATTTGCTG 96755  
Qy 300 sGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAl 320  
Db 96754 TGAGCGCAGTGGGTGGAGTACCGCTGTGTCAGGCTGCTAGCTGTCTCAAAAGCGGGC 96695  
Qy 320 aValLeuGlyArgThrLeuGlyValHisProAspAspAspLeuLeuPheThrValPheSe 340  
Db 96694 CTTGCTTGGCAGGACCTTGGAGTCCATCCAGATGATGAGCTCTCTTACCGCTTCTC 96635  
Qy 340 rLysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysLlePheLleLe 360  
Db 96634 CAAGGCGCAGAGCGGAAATGAATCCCTGGATGAGTCGCGCTGTGTCTCTTCACTCT 96575  
Qy 360 uLysGlnLleAsnAspArgLleLysGluArgLeuGlnSerCystyrArgGlyGluGlyTh 380  
Db 96574 GAAGCAGATTAATGACCGCATTAAGGAGCGGCTGACGTCTGTGTACCGGGCGAGGCAC 96515  
Qy 380 rLeuAspLeuAlaTrpLeuLysValLysAspLleProCysSerSerAla 396  
Db 96514 GCTGGACCTGGCTTGGCTCAAGCTGAAGGACATCCCTCGCAGCAGTSCG 96466  
RESULT 8  
HSM802037 1029 bp mRNA linear PRI 18-FEB-2000  
LOCUS Homo sapiens mRNA; cDNA DKFp434G0625 (from clone DKFp434G0625);  
DEFINITION partial cds.  
ACCESSION AL137352  
VERSION AL137352.1 GI:6807867  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1029)  
AUTHORS Ansoerge.W., Wirkner,U., Mewes,H.W., Gassenhuber,J. and Wiemann,S.  
TITLE Direct Submission  
JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitze 18a, D-82152  
Martinsried, GERMANY  
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by EMBL (European Molecular Biology Laboratories,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFp434G0625) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further  
information about the clone and the sequencing project is available  
at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES  
Location/Qualifiers  
source  
1..1029

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434G0625"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..955
/gene="DKFZp434G0625"
<1..955
/gene="DKFZp434G0625"
/note="strong similarity to plexins"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAB70707.1"
/db_xref="GI:6807868"
/translation="KNSKADGMEFAYVFHDFEFAVMKIPSDTEIIPDEDIYVYVGF
SGNFVFLIQEMVSPGSKTEQYVTSKLVRLCKREDTAFNSIVEVPGICRSRGVEY
RLQAAYLSKAGAVLGRITLVHPDDDLFTVFSKQKRMKSLDESALCIFILKQIND
RIKERQSCYRGEGTLDLAWLKVDIPCSALITIDNFCGLDMNAPLGVSDMVRGIP
VFTEDRDMTIVAYVYKXHSIAFVGTGKGLKKSFTGPGGQITQEWIGVEGDPPGA
NIASQBMCLVVLQCSSHKAISDRVQPLLCFLNVPNGSS"
polyA_signal
991..996
polyA_site
1006
BASE COUNT 258 a 260 c 264 g 247 t
ORIGIN
Alignment Scores:
Pred. No.: 6 09e-268 Length: 1029
Score: 252.00 Matches: 252
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.29% Indels: 0
DB: 9 Gaps: 0
US-09-964-956-13 (1-1896) x HSM802037 (1-1029)
QY 206 LysAsnSerGluAlaAspGlyMetPheAlaTyrrValPheHisAspGluPheValAlaSer 225
Db 2 AAGAACTCTGAGCGGATGCATGTTTCGCGTACGCTTCCATGATGATGATTCGTCGCGCTCG 61
QY 226 MetIleLysIleProSerAspThrPheThrIleIleProAspPheAspIleTyrrVal 245
Db 62 ATGATTAAGATCCCTTCGGACACCTTCACCATCATCCCTGACCTTGATATCTACTATGTC 121
QY 246 TyrGlyPheSerGlyAsnPheValTyrrPheLeuThrLeuGlnProGluMetValSer 265
Db 122 TATGGTTTATGACAGTGGCACTTTGCTACTTTTGGACCTCAACCTGAGATGGTGCT 181
QY 266 ProGlySerThrThrLysGluGlnValTyrrThrSerLysLeuValArgLeuCysLys 285
Db 182 CCACCAAGGCTCCACCAAGGAGCAGGTGTATACATCCCAAGCTCGTGAGGCTTTGCAAG 241
QY 286 GluAspThrAlaPheAsnSerTyrrValGluValProIleGlyCysGluArgSerGlyVal 305
Db 242 GAGGACACAGCCTTCAACCTCTATGATAGAGTGCCCATTTGGCTGTAGCCAGCTGGGGTG 301
QY 306 GluTyrrArgLeuLeuGlnAlaAlaTyrrLeuSerLysAlaGlyAlaValLeuGlyArgThr 325
Db 302 GAGTACCGCTGCTGACAGCTGCTACCTGCTCCAAAGCGGGGCGGCTGTGTCAGAGACC 361
QY 326 LeuGlyValHisProAspAspLeuLeuPheThrValPheSerLysGlyGlnLysArg 345
Db 362 CTGGAGTCCATCCAGATGATGACCTGCTCTTCCAGCTCTTCCAAAGGCGGCAAGACCG 421
QY 346 LysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysGlnIleAsnAsp 365
Db 422 AAAATGAAATCCCTGGATGAGTCGGCCCTGTGTCATCTTCAATGAGCAGATAAATGAC 481
QY 366 ArgIleLysGluArgLeuGlnSerCysTyrrArgGlyGluGlyThrLeuAspLeuAlaTrp 385
Db 482 CGCATTAAGAGCGGCTGCAGTCTTGTACCGGGGCGAGGCGGACCTGGACCTGGCGCTGG 541
```

```
386 LeuLysValLysAspIleProCysSerAlaLeuLeuThrIleAspAsnPhcCys 405
Db 542 CTTAAGGTGAAGGACATCCCTCGCAGCAGTCGGCTCTTAACCATTTGACGATAACTTCTGT 601
QY 406 GlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArgGlyIleProValPhe 425
Db 602 GGCCTGGACATGAATGCTCCCTCGGAGTCCCGCATGGTGGTGGATTCCTCCGCTCTTC 661
QY 426 ThrGluAspArgAspArgMetThrSerValIleAlaTyrrValTyrrLysAsnHisSerLeu 445
Db 662 ACGGAGGACAGGACCGCATGCTGCTCATGCTATGCTACAGAACCACTCTCTG 721
QY 446 AlaPheValGlyThrLysSerGlyLysLeuLysLys 457
Db 722 GCCTTTGTGGCCACCAAAAGTGGCAAGCTGAAGAG 757
RESULT 9
AC025595/3
LOCUS
DEFINITION Homo sapiens clone RP11-24E10, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC025595
VERSION AC025595.2 GI:7341984
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 74655)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-24E10
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 74655)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,N., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye.W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:7229794.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4440
Center clone name: 24_E10
-----
* NOTE: This record contains 90 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
```

\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

\* 727: contig of 727 bp in length  
\* 728 827: gap of 100 bp  
\* 828 1590: contig of 763 bp in length  
\* 1591 1690: gap of 100 bp  
\* 1691 2443: contig of 753 bp in length  
\* 2444 2543: gap of 100 bp  
\* 2544 3286: contig of 743 bp in length  
\* 3287 3386: gap of 100 bp  
\* 3387 4099: contig of 713 bp in length  
\* 4100 4199: gap of 100 bp  
\* 4200 4924: contig of 725 bp in length  
\* 4925 5024: gap of 100 bp  
\* 5025 5745: contig of 721 bp in length  
\* 5746 5845: gap of 100 bp  
\* 5846 6567: contig of 722 bp in length  
\* 6568 6667: gap of 100 bp  
\* 6668 7409: contig of 742 bp in length  
\* 7410 7509: gap of 100 bp  
\* 7510 8239: contig of 730 bp in length  
\* 8240 8339: gap of 100 bp  
\* 8340 9048: contig of 709 bp in length  
\* 9049 9148: gap of 100 bp  
\* 9149 9902: contig of 754 bp in length  
\* 9903 10002: gap of 100 bp  
\* 10003 10745: contig of 743 bp in length  
\* 10746 10845: gap of 100 bp  
\* 10846 11593: contig of 748 bp in length  
\* 11594 11693: gap of 100 bp  
\* 11694 12426: contig of 733 bp in length  
\* 12427 12526: gap of 100 bp  
\* 12527 13240: contig of 714 bp in length  
\* 13241 13340: gap of 100 bp  
\* 13341 14074: contig of 734 bp in length  
\* 14075 14174: gap of 100 bp  
\* 14175 14906: contig of 732 bp in length  
\* 14907 15006: gap of 100 bp  
\* 15007 15728: contig of 722 bp in length  
\* 15729 15828: gap of 100 bp  
\* 15829 16542: contig of 714 bp in length  
\* 16543 16642: gap of 100 bp  
\* 16643 17379: contig of 737 bp in length  
\* 17380 17479: gap of 100 bp  
\* 17480 18225: contig of 746 bp in length  
\* 18226 18325: gap of 100 bp  
\* 18326 19056: contig of 731 bp in length  
\* 19057 19156: gap of 100 bp  
\* 19157 19894: contig of 738 bp in length  
\* 19895 19994: gap of 100 bp  
\* 19995 20729: contig of 735 bp in length  
\* 20730 20829: gap of 100 bp  
\* 20830 21569: contig of 740 bp in length  
\* 21570 21669: gap of 100 bp  
\* 21670 22401: contig of 732 bp in length  
\* 22402 22501: gap of 100 bp  
\* 22502 23225: contig of 724 bp in length  
\* 23226 23325: gap of 100 bp  
\* 23326 24053: contig of 728 bp in length  
\* 24054 24153: gap of 100 bp  
\* 24154 24892: contig of 739 bp in length  
\* 24893 24992: gap of 100 bp  
\* 24993 25727: contig of 735 bp in length  
\* 25728 25827: gap of 100 bp  
\* 25828 26566: contig of 739 bp in length  
\* 26567 26666: gap of 100 bp  
\* 26667 27410: contig of 744 bp in length  
\* 27411 27510: gap of 100 bp

\* 27511 28241: contig of 731 bp in length  
\* 28242 28341: gap of 100 bp  
\* 28342 29090: contig of 749 bp in length  
\* 29091 29190: gap of 100 bp  
\* 29191 29923: contig of 733 bp in length  
\* 29924 30023: gap of 100 bp  
\* 30024 30742: contig of 719 bp in length  
\* 30743 30842: gap of 100 bp  
\* 30843 31558: contig of 716 bp in length  
\* 31559 31658: gap of 100 bp  
\* 31659 32388: contig of 730 bp in length  
\* 32389 32488: gap of 100 bp  
\* 32489 33204: contig of 716 bp in length  
\* 33205 33304: gap of 100 bp  
\* 33305 34018: contig of 714 bp in length  
\* 34019 34118: gap of 100 bp  
\* 34119 34850: contig of 732 bp in length  
\* 34851 34950: gap of 100 bp  
\* 34951 35689: contig of 739 bp in length  
\* 35690 35789: gap of 100 bp  
\* 35790 36536: contig of 747 bp in length  
\* 36537 36636: gap of 100 bp  
\* 36637 37354: contig of 718 bp in length  
\* 37355 37454: gap of 100 bp  
\* 37455 38198: contig of 744 bp in length  
\* 38199 38298: gap of 100 bp  
\* 38299 39045: contig of 747 bp in length  
\* 39046 39145: gap of 100 bp  
\* 39146 39872: contig of 727 bp in length  
\* 39873 39972: gap of 100 bp  
\* 39973 40679: contig of 707 bp in length  
\* 40680 40779: gap of 100 bp  
\* 40780 41511: contig of 732 bp in length  
\* 41512 41611: gap of 100 bp  
\* 41612 42342: contig of 731 bp in length  
\* 42343 42442: gap of 100 bp  
\* 42443 43178: contig of 736 bp in length  
\* 43179 43278: gap of 100 bp  
\* 43279 44022: contig of 744 bp in length  
\* 44023 44122: gap of 100 bp  
\* 44123 44860: contig of 738 bp in length  
\* 44861 44960: gap of 100 bp  
\* 44961 45676: contig of 716 bp in length  
\* 45677 45776: gap of 100 bp  
\* 45777 46515: contig of 739 bp in length  
\* 46516 46615: gap of 100 bp  
\* 46616 47350: contig of 735 bp in length  
\* 47351 47450: gap of 100 bp  
\* 47451 48148: contig of 698 bp in length  
\* 48149 48248: gap of 100 bp  
\* 48249 48949: contig of 701 bp in length  
\* 48950 49049: gap of 100 bp  
\* 49050 49775: contig of 726 bp in length  
\* 49776 49875: gap of 100 bp  
\* 49876 50592: contig of 717 bp in length  
\* 50593 50692: gap of 100 bp  
\* 50693 51419: contig of 727 bp in length  
\* 51420 51519: gap of 100 bp  
\* 51520 52272: contig of 753 bp in length  
\* 52273 52372: gap of 100 bp  
\* 52373 53123: contig of 751 bp in length  
\* 53124 53223: gap of 100 bp  
\* 53224 53965: contig of 742 bp in length  
\* 53966 54065: gap of 100 bp  
\* 54066 54780: contig of 715 bp in length  
\* 54781 54880: gap of 100 bp  
\* 54881 55621: contig of 741 bp in length  
\* 55622 55721: gap of 100 bp  
\* 55722 56458: contig of 737 bp in length  
\* 56459 56558: gap of 100 bp  
\* 56559 57291: contig of 733 bp in length  
\* 57292 57391: gap of 100 bp

```
Alignment Scores:
Pred. No.: 6.03e-211 Length: 74655
Score: 202.00 Matches: 202
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.65% Indels: 0
DB: 2 Gaps: 0

US-09-964-956-13 (1-1896) x AC025595 (1-74655)
QY 17 MetValGlyMetGlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLys 36
Db 34015 ATGGTGGCATGGCTCTCCACTTGTCTCACCAGGAGGAGCCGCTGTCACCAAG 33956
QY 37 GlnArgSerPheValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValVal 56
Db 33955 CACGGTCATTTGCACATTCGAGGAGAGCCGCCGAGGGTTTCATCACCCTGGTGTG 33896
QY 57 AspGluArgThrGlyHisLleTyrLeuGlyAlaValAsnArgLleTyrLysLeuSerSer 76
Db 33895 GATGAGAGGACAGACACATTTACTTGGGGCGCTCAATCGGATTTTACAAGCTCTCCAGC 33836
QY 77 AspLeuLysValLeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyr 96
Db 33835 GACCTGAAGGCTCTGGTGACGCATGAGACAGGCGCGGACGAGACACCCCAAGTGTTC 33776
QY 97 ProProArgLleValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLys 116
Db 33775 CCACCCGCATGCTCCAGACCTGCAATGAGCCCTGACCCACCATGTCACCAAG 33716
QY 117 MetLeuLeuLleAspTyrLysGluAsnArgLeuLeuAlaCysGlySerLeuTyrGlnGly 136
Db 33715 ATGCTCTCATAGACTACAAGGAGAACAGGCTGATTCGCTGTGGAGCCTGTACCAAGGC 33656
QY 137 IleCysLysLeuLeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLys 156
Db 33655 ATCTGAAGCTGCTGAGCTGGAGGACCTTCAAGCTGGGGAGCCCTTATCATAAAG 33596
QY 157 GluHisTyrLeuSerGlyValAsnGluSerGlySerValPheGlyValIleValSerTyr 176
Db 33595 GAGCACTATCTGTCAGGTGTCACGAGAGCGCTCGGTCTTTGGAGTGATCGTCTCTAC 33536
QY 177 SerAsnLeuAspAspLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPhe 196
Db 33535 AGCAACCTGGATGACAGAGCTGTTTCATTGCCACGCGAGTGGTGGAAAGCCCGAGTATTT 33476
QY 197 ProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyr 216
Db 33475 CCACCATCTCCAGCGGAAACTGACCAAGAACTCTGAGCGGATGTCATGTCGCGTAC 33416
QY 217 ValPhe 218
Db 33415 GTCTTT 33410

RESULT 10
HSM805617
LOCUS Homo sapiens mRNA 3556 bp linear PRI 12-JUL-2002
DEFINITION Homo sapiens mRNA; cDNA DKFzp56600546 (from clone DKFzp56600546).
ACCESSION AL834504
VERSION AL834504.1 GI:21740304
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3556)
AUTHORS Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuberberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
```

sequencing consortium of the German Genome Project. This clone (DKFzp56600546) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES  
Location/Qualifiers  
1..3556  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFzp56600546"  
/tissue\_type="kidney"  
/clone\_lib="566 (synonym: hfkcd2). Vector pAMP1; host  
XL2blue; sites NotI + SalI"  
/dev\_stage="fetal"  
1..3556  
/gene="DKFzp56600546"  
/cDS  
1..1031  
/gene="DKFzp56600546"  
/note="strong similarity to mouse plexin 2"  
/codon\_start=3  
/product="hypothetical protein"  
/protein\_id="CAD39161.1"  
/db\_xref="GI:21740305"  
/translation="LEWQGGSGARMILQEDITTKIENDWKRNLTLAHYQVDPGYSVA  
LVSKQVTAYNAVNNSTVSRTSASKYENIRVCTSPDSLSRSPMTITPDLESQVKKMHL  
VKNHEHGDKGDRGSKMVSEIYLTRLATKTGLOKFVDDLFTETIFSTAHRSALPLA  
IKYMFDFLEQADKHGIDHPVHTWKNCLPLRFWDMKNPQVDFDIKNSITDAC  
LSVVAOTFMDSCSTSEHRLGDKSPSKLLIADKVFYKNWERYTSDICKMPSIDQD  
MNLAEOSRHMNEFNTMSALSEIFSYGVKYSSEILGLDHDHDDQCGKOKLAYKLEQV  
ITLMSLDS"  
polyA\_site  
3532  
/gene="DKFzp56600546"  
BASE COUNT 927 a 907 c 892 g 830 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.27e-206 Length: 3556  
Score: 197.00 Matches: 340  
Percent Similarity: 98.84% Conservative: 0  
Best Local Similarity: 98.84% Mismatches: 2  
Query Match: 10.39% Indels: 4  
DB: 9 Gaps: 0  
US-09-964-956-13 (1-1896) x HSM805617 (1-3556)  
QY 1553 LeuGluTyrArgGlnGlySerGlyAlaArgMetIleLeuGlnAspLysPheThrThr 1572  
Db 3 CTGGAGTGGCGACAAAGGAAGTGGGCAAGGATGATCTTCGAGGATGAACATCACACC 62  
QY 1573 LysIleGluAsnAspTyrLysArgLeuAsnThrLeuAlaHisTyrGlnValProAspGly 1592  
Db 63 AAGATTGAGATGATTGGGAAGGAGCTGAACACACTGGCCACTACCAGGTGCAGATGGT 122  
QY 1593 SerValValAlaLeuValSerLysGlnValThrAlaTyrAsnAlaValAsnAsnSerThr 1612  
Db 123 TCCGTGGTGGCATTAGTGTCCAAGCAGGTGACAGCTATAAGCAGTGAACAACTCCACC 182  
QY 1613 ValSerArgThrSerAlaSerLysTyrGluAsnMetIleArgTyrThrGlySerProAsp 1632  
Db 183 GTCTCCAGGACCTCAGCAAGTAAATATGAAACATGATCCGGTACACGGGAGCCCGGAC 242  
QY 1633 SerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyValLysMetTyr 1652  
Db 243 AGCCTCCCTCAGGACACCTATGATGATCACTCTGACCTGGAGAGTGAAGTCAAGATGTGG 302  
QY 1653 HisLeuValLysAsnHisGluIleHisGlyAspGlnLysGluGlyAspArgGlySerLysMet 1672  
Db 303 CACCTAGTGAAGAACACGAGCAGCAGACGAGACCAAGAGGAGGAGCCGGGAGCAAGATG 362  
QY 1673 ValSerGluIleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheVal 1692  
Db 363 GTGTCTGAAATCTACCTGACCCGACTCCTGGCCACTAAGGGGCACACTGCAGAAAGTTTGTG 422

Percent Similarity:	98.99%	Conservative:	0
Best Local Similarity:	98.99%	Mismatches:	1
Query Match:	6.54%	Indels:	2
DB:	6	Gaps:	0
US-09-964-956-13 (1-1896) x:AX088116 (1-601)			
QY	503	ThrArgValProValGluSerCysGlyGlnTyrGlnSerCysGlyGluCysLeuGlySer	522
DB	2	ACCAGAGTCCCTGTGGAGTCTGTGGTCAATATCAGAGCTGGCGGAGTGGCTTGGCTCA	61
QY	523	GlyAspProHisCysGlyTrpCysValLeuHisAsnThrCysThrArgLysGluArgCys	542
DB	62	GGCGACCCCACTGTGGTGGTGTGGTGCACACACTTGCACCGAAGGAGCGGTGT	121
QY	543	GluArgSerLysGluProArgPheAlaSerGluMetLysGlnCysValArgLeuThr	562
DB	122	GAGCGGTCCAGAGCGCCCGAGTTGGCTCGGAGATGAGCAGTGTGTCGCGTGCAGC	181
QY	563	ValHisProAsnAsnTleSerValSerGlnTyrAsnVal-LeuLeuValLeuGluThrTy	582
DB	182	GTCCATCCCAACAATATCTCCGCTCTCAGTACACAGC-GCTGCTGGTCTCGAGAGTA	240
QY	582	rAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSerGluMetAs	602
DB	241	CAATGTCGCGGAGCTGTACGTCGCTCACTCACTTTCAGGACCTGTTCAGAGATGGA	300
QY	602	pGlyLeuValValGlyAsnGlnTleGlnCysTyrSerProAlaAlaLysGluValProAr	622
DB	301	TGGGCTGGTGGTGGCAATCAGATCCAGTCCAGTCTACTCCCTGCAGCCAGGAGTGC	360
QY	622	gilelleThrGluAsnGlyAspHisValValGlnLeuGlnLeuLysSerLysGluTh	642
DB	361	GATCATCACAGAAATGGGACCACTGTCGTCAGCTTCAGCTCAATCAAGAGAGAC	420
QY	642	rGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerValHisAsnSerCy	662
DB	421	CGCATGACCTTCGCCAGCACCACTTGTCTTCTACAAATGCAGCGTCCACAAATTCGTG	480
QY	662	sLeuSerCysValGluSerProTyrArgCysHisTrpCysLysTyrArgHisValCysTh	682
DB	481	CCTGTCTCTGGTGGAGTCCATACCGCTGCCACTGGTGTAAATACCGCATGTCTGCAC	540
QY	682	rHisAspProLysThrCysSerPheGlnGluGlyArgValLysLeuProGlu	699
DB	541	CCATGACCCCAACACCTGCTCTTCAGGAGCCGCGAGTGAAGCTGCCCGAG	592
RESULT 12			
AK091538	3560 bp mRNA linear PRI 15-JUL-2002		
LOCUS	Homo sapiens CDNA FLJ34219 fls, clone FCBBF3022321, moderately		
DEFINITION	similar to PLEXIN 4 PRECURSOR.		
ACCESSION	AK091538		
VERSION	AK091538.1 GI:21749932		
KEYWORDS	oligo capping; fls (full insert sequence).		
SOURCE	Homo sapiens fetal brain cDNA to mRNA, clone_lib:FCBBF3		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S., Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M., Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Matsumoto,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masubo,Y., Nagai,K. and Isogai,T.		
TITLE	NEO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3560)		

QY	1693	AspAspLeuPheGluThrIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAla	1712
DB	423	GATGACCTCTTTGAGACCATCTTCAGCAGCGCACCGTGGCTCTGCGCTGGCC	482
QY	1713	IleLysTyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspPro	1732
DB	483	ATCAAGTACATGTTGACTTCTCGATGAGCAGGCTGATAACATGGCATTCATGACCCG	542
QY	1733	HisValArgHisThrTrpLysSerAsnCysLeuProLeuArgPheTrpValAsn-MetII	1752
DB	543	CACGTCGCCCATACCTGGAAGAGAAATTCGCTGCCCTGAGGTTTGGGTGCA-CATGAT	601
QY	1752	eLysAsnProGlnPheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSe	1772
DB	602	CAAGAACCCGAGTTTGCTTTGACATCCATTAAGACAGCATCAGACGCTGCTCTC	661
QY	1772	rValValAlaGlnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAs	1792
DB	662	TGTGTGGCTCAGACCTTCATGGACTCTTGCTCCAGTCCAGACCGGCTGGGCAAGGA	721
QY	1792	pSerProSerAsnLysLeuLeuTyrAlaLysAspIle-ProSerTyrLysAsnTrpValG	1812
DB	722	CTCGCCCTCCCAACAGCTGCTATGCCCAGAGACCT-CCCAAGCTACAGAAATGGGTGG	780
QY	1812	luArgTyrTyrSerAspIleGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaT	1832
DB	781	AGAGTATTACTCAGACATAGGAAGATGCCAGCCATCAGCGCAACAGCATGAACGCAT	840
QY	1832	yrLeuAlaGluGlnSerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerG	1852
DB	841	ACCTGGCTGACGAGTCCCGGATGTCATGATGAGTTCACACCATGAGTGCACCTCTAG	900
QY	1852	luIlePheSerTyrValGlyLysTyrSerGluGluIleLeuGlyProLeuAspHisAspA	1872
DB	901	AGATCTTCTCTATGTGGCAATACAGCGAGAGATCCTTGACCTCTGGACCAAGGAGC	960
QY	1872	spGlnCysGlyLysGlnLysLeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerL	1892
DB	961	ACCAGTGTGGGAAGCAGAACTGGCTACAAACTAGAACAGTCAACCTCATGAGCT	1020
QY	1892	euAspSer	1894
DB	1021	TAGACAGC	1028
RESULT 11			
AX088116	601 bp DNA linear PAT 17-MAR-2001		
LOCUS	Sequence 7 from Patent WO0114420.		
DEFINITION	AX088116		
ACCESSION	AX088116.1 GI:13397033		
VERSION	1 (bases 1 to 601)		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 601)		
AUTHORS	Artigiani,S., Comoglio,P.M., Goodman,C.S., Tesier-Lavigne,M.R. and Tamagnone,L.		
TITLE	Novel plexins and uses thereof		
JOURNAL	Patent: WO 0114420-A 7 01-MAR-2001.		
FEATURES	University of Torino (IT) ; The Regents of the University of California (US)		
Source	Location/Qualifiers		
	1..601		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	124 a	182 c	173 g 122 t
ORIGIN			
Alignment Scores:	1.36e-126	Length:	601
Pred. No.:	124.00	Matches:	196
Score:			

AUTHORS Isogai, T. and Yamamoto, J.  
TITLE Direct Submission  
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 202-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.

## FEATURES

source

1. .3560  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="FCBBF3022321"  
/tissue\_type="brain"  
/clone\_lib="FCBBF3"  
/dev\_stage="fetal"  
/note="cloning vector: pME18SFL3"

BASE COUNT 901 a 881 c 877 g 901 t

## ORIGIN

Alignment Scores:

Pred. No.: 1.13e-123 Length: 3560  
Score: 122.00 Matches: 208  
Percent Similarity: 99.05% Conservatives: 0  
Best Local Similarity: 99.05% Mismatches: 1  
Query Match: 6.43% Indels: 2  
DB: Gaps: 0

US-09-964-956-13 (1-1896) x AK091538 (1-3560)

QY 1686 GlyThrLeuGlnLysPheValAspLeuPheGluThrIlePheSerThrAlaHisArg 1705  
|||||  
Db 51 GGCACACTGCAGAAAGTTGGGATGACCTTTTGAGACCATTTCACACGGCACACCGT 110  
|||||  
QY 1706 GlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAlaAsp 1725  
|||||  
Db 111 GGCTGCTCCCTGCCCTGGCCATCAAGTACATGTTTGACTTCCTGGATGACGCGTGAT 170  
|||||  
QY 1726 LysHisGlyIleHisAspProHisValArgHisThrTrpLysSerAsnCysLeuProLeu 1745  
|||||  
Db 171 AAACATGGCATTCATACCCGCGCAGCTGGCCATACCTGGGAAGACAAATTCCTGCCCTG 230  
|||||  
QY 1746 ArgPheTrpValAsnMetIleLysAsnProGlnPheValPheAspIleHisLysAsnSer 1765  
|||||  
Db 231 AGTTTGGGTCAACATGATCAAGAACCCCGCAGTTTGTGTGACATCCATAAGACAGC. 290  
|||||  
QY 1766 IleThrAspAlaCysLeuSerValValAlaGlnThrPheMetAspSerCysSerThrSer 1785  
|||||  
Db 291 ATCAGACAGCCCTGCTCTCTGTGGTGCCTCAGACCTTCATGACTCTTCTCCACGTC 350  
|||||  
QY 1786 GluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysAspIlePro 1805  
|||||  
Db 351 GAGCAGCGGTGGCGAAGGACTCGCCCTCCCAAGCTGCTGTATGCGCAAGGACATCCCC 410  
|||||  
QY 1806 SerTyrLysAsnTrpValGluArgTyrTyrSerAspIleGlyLysMetProAlaIleSe 1825  
|||||  
Db 411 AGCTACAG-GAATTTGGTGGAGAGGTATTACTCAGACATAGGGAAGATGCGACCATCAG 469  
|||||  
QY 1825 rAspGlnAspMetAsnAlaTyrLeuAlaGluGlnSerArgMetHisMetAsnGluPheAs 1845  
|||||  
Db 470 CGACCAAGACATCAGCATACCTGGCTGAGCAGTCCGCTATGCACATGATGAGTCA 529  
|||||  
QY 1845 nThrMetSerAlaLeuSerGluIlePheSerTyrValGlyLysTyrSerGluGluIleLe 1865  
|||||  
Db 530 CACCATGAGTGCACTCTCAGACATCTTCTCTATGTGGCAAAATACAGCAGGAGATCCT 589  
|||||  
QY 1865 uGlyProLeuAspHisAspGlnCysGlyLysGlnLysLeuAlaTyrLysLeuGluG 1885  
|||||

Db 590 TGGACCTCTGGACCACGACGACCGACCTGTGGGAAGCAGAAACTGGCCTACAAACTAGACA 649  
QY 1895 nValIleThrLeuMetSerLeuAspSer 1894  
|||||  
Db 650 AGTCATAACCCCTCATGAGCTTAGACAGC 677  
RESULT 13  
AC131554/c 105810 bp DNA linear HTG 24-AUG-2002  
LOCUS Rattus norvegicus clone CH230-10113, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 46 unordered pieces.  
AC131554  
VERSION AC131554.1 GI:22474821  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 105810)

## REFERENCE

## AUTHORS

Muzny D.Marie., Metzker M.Lee., Abramson S., Adams C., Alder J.,  
Allen C., Allen H., Alsbrooks S., Amin A., Anguiano D.,  
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,  
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,  
Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,  
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,  
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,  
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,  
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,  
Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,  
Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K.,  
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,  
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,  
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,  
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,  
Gebregorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W.,  
Gunaratne P., Healand W., Hamill C., Hamilton C., Hamilton K.,  
Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,  
Hernandez R., Hines S., Hladun S.L., Hodgson A., Hoques M.,  
Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,  
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,  
Karpathy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,  
Kowls C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,  
Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,  
Lorensheewa L., Loulseghe H., Lozado R.J., Lu X., Ma J.,  
Maheshwari M., Mahindartne M., Mahmoud M., Malloy K., Mangum A.,  
Mangum B., Mapa P., Martin K., Martin R., Martinez E.,  
Mawhinney S., McLeod M., McNeill T., Meenen E., Milosavljevic A.,  
Miner G., Minja E., Montemayor J., Moore S., Morgan M., Morris K.,  
Morris S., Munidasa M., Murphy M., Nair L., Nankervis C., Neal D.,  
Newton N., Nguyen N., Norris S., Nwaokemele O., Okwuonu G.,  
Olarnunsaqoon A., Pal S., Parks K., Pasternak S., Paul H.,  
Perez A., Perez L., Pfannkuch C., Plopper F., Poindexter A.,  
Popovic D., Primus E., Pu L.-L., Puzos M., Quiroz J., Rachlin E.,  
Reeves K., Regier M.A., Reigh R., Reilly B., Reilly M., Ren Y.,  
Reuter M., Richards S., Riggs F., Rives C., Rodkey T., Rojas A.,  
Rose M., Rose R., Ruiz S.J., Sanders W., Savery G., Scherer S.,  
Scott G., Shatsman S., Shen H., Shetty J., Shvartsbeyn A.,  
Sison I., Sitter C.D., Smajd J., Sneed A., Sodergren E.,  
Song X.-Z., Sorelle R., Sosa J., Steimle M., Strong R., Sutton A.,  
Svatek A., Tabak P., Taylor C., Taylor T., Thomas N., Thomas S.,  
Tingey A., Trejos Z., Usmani K., Valas R., Vera V., Villasana D.,  
Waldron L., Walker B., Wang J., Wang Q., Wang S., Warren J.,  
Warren R., Wei X., White F., Williams G., Willson R., Wleczyk R.,  
Wooden H., Worley K., Wright D., Wright R., Wu J., Yakub S.,  
Yen J., Yoon L., Yoon V., Yu F., Zhang J., Zhou J., Zhou X.,  
Zhao S., Dunn D., von Niederhausern A., Weiss R., Smith D.R.,  
Holt R.A., Smith H.O., Weinstein G. and Gibbs R.A.

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

Unpublished  
Direct Submission

2 (bases 1 to 105810)

Rat Genome Sequencing Consortium.

Direct Submission





\* 3266 4837: contig of 1572 bp in length  
\* 4838 4937: gap of 100 bp  
\* 4938 45186: contig of 40249 bp in length  
\* 45187 45285: gap of 100 bp  
\* 45287 57434: contig of 12148 bp in length  
\* 57435 57534: gap of 100 bp  
\* 57535 84630: contig of 27096 bp in length  
\* 84631 84730: gap of 100 bp  
\* 84731 111275: contig of 26545 bp in length  
\* 111276 111375: gap of 100 bp  
\* 111376 134525: contig of 23150 bp in length  
\* 134526 134625: gap of 100 bp  
\* 134626 170655: contig of 36030 bp in length  
\* 170656 170755: gap of 100 bp  
\* 170756 173385: contig of 2630 bp in length.

FEATURES

source  
1. .173385  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-462013"  
/clone\_lib="RPC1-23 Female Mouse BAC"  
misc\_feature  
1. 715  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
816. .1666  
/note="assembly\_fragment"  
misc\_feature  
1767. 3165  
/note="assembly\_fragment"  
misc\_feature  
3266. .4837  
/note="assembly\_fragment"  
misc\_feature  
4938. .45186  
/note="assembly\_fragment"  
misc\_feature  
45287. 57434  
/note="assembly\_fragment"  
misc\_feature  
57535. .84630  
/note="assembly\_fragment"  
misc\_feature  
84731. .111275  
/note="assembly\_fragment"  
misc\_feature  
111376. .134525  
/note="assembly\_fragment"  
misc\_feature  
134626. .170655  
/note="assembly\_fragment"  
misc\_feature  
170756. .173385  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right  
BASE COUNT 47462 a 38869 c 40100 g 45911 t 1043 others  
ORIGIN

Alignment Scores:  
Pred. No.: 9,61e-87 Length: 173385  
Score: 90.00 Matches: 90  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.75% Indels: 0  
DB: 2 Gaps: 0

US-09-964-956-13 (1-1896) x AC102171 (1-173385)

QY 1340 ValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIle 1359  
DB 35099 GTCCAGGTCACCGCAGGCGGTGGAGAAAGCACTCAAGCTCTTCGCCAGGTCATC 35158  
QY 1360 AsnAsnLysValPheLeuLeuSerPheTleArgThrLeuGluSerGlnArgSerPheSer 1379  
DB 35159 AACACACAGGTGTCCTCTCTTTCATCCGCACTCTGGAGTCTCAGCGCAAGTTTCC 35218  
QY 1380 MetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeuGlu 1399  
DB 35219 ATGCGCAGCCAGGTAAGCGTGCCCTCCCTCATCATGACGGTACTTCAGACCAAGCTAG 35278  
QY 1400 TyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGluSer 1419

Db 35279 TATGCCACTGAGCTGCTGAAGCAGCTACTGGTGCTACCTCATTGATGAAGACCTGGAGAGC 35338  
QY 1420 LysAsnHisProLysLeuLeuLeuArgArg 1429  
Db 35339 AAGAACCAACCCCAAGCTGCTCTCAGGAGG 35368  
RESULT 15  
AC018643/c 198470 bp DNA linear PRI 17-NOV-2000  
LOCUS Human Chromosome 7 clone RP11-297N5, complete sequence.  
DEFINITION AC018643  
AC018643.3 GI:95586608  
VERSION AC018643.3  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 198470)  
Kaul,R.K., Zhou,Y., James,R.A., Raymond,C., Haugen,E.D. and  
Olson,M.V.  
Large-scale Mapping and Sequencing of Human Chromosome 7  
Unpublished  
2 (bases 1 to 198470)  
Hubb,K.L., Desmarais,C.L., Ramsey,S.A. and Hubley,R.M.  
Direct Submission  
Submitted (15-DEC-1999) Human Genome Center, University of  
Washington, Box 352145, Seattle, WA 98195, USA  
3 (bases 1 to 198470)  
Kaul,R.K., Zhou,Y., James,R.A., Raymond,C., Haugen,E.D. and  
Olson,M.V.  
Direct Submission  
Submitted (28-JUL-2000) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
4 (bases 1 to 198470)  
Kaul,R.K., Zhou,Y., James,R.A., Raymond,C., Haugen,E.D. and  
Olson,M.V.  
Direct Submission  
Submitted (17-NOV-2000) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Jul 28, 2000 this sequence version replaced gi:8050921.  
----- Genome Center  
Center: University of Washington Genome Center  
Center Code: UWGC  
Web site: http://www.genome.washington.edu/UWGC/  
Contact: uwgchgs@u.washington.edu  
----- Project Information  
Center project name: HsaChr7  
Center clone name: RP11-297N5 (djs233)  
----- Summary Statistics  
Sequencing vector: M13; X02513; 100% of reads  
Chemistry: Dye-primer-amersham; 99% of reads  
Chemistry: Dye-terminator ET-amersham; 1% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 198174 bases at least Q40  
Consensus quality: 198402 bases at least Q30  
Consensus quality: 198464 bases at least Q20  
Insert size: 212411; 7.1% error; agarose-fp  
Quality coverage: 7.50x in Q20 bases; agarose-fp  
Quality coverage: 8.02x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP11-533N5 AC055877  
3': RP11-348A15 (UWGC:djs302) AC009364, 64369-bp clone overlap

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

# Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

ECORI NsiI BglII

SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt
8696	8806	14309	14401	3995	4156
2842	2841	1127	1148	2067	2070
2747	2713	1080	1148	4799	4913
8210	8230	6772	6747	2868	2861
11559	11374	667	<800	22606	22935
3869	3919	15	<800	1547	1553
792	795	3544	3491	924	925
10723	10666	1175	1148	13450	12996
2047	2049	1167	1148	5240	5230
2240	2261	386	<800	1342	1322
6878	6811	2690	2607	8	<800
923	920	4735	4641	71	<800
6549	6516	89	<800	11499	11149
4450	4467	4652	4641	435	<800
534	<800	3773	3781	4540	4701
638	<800	3018	2967	7275	7301
13772	13792	2040	2065	1104	1090
12138	12156	273	<800	4304	4539
6454	6516	3873	3854	10	<800
3091	3116	8082	7991	14	<800
91	<800	1210	1173	4144	3904
2427	2388	4800	4791	679	<800
2058	2049	2919	2967	794	774

3598	3609	333	<800	3761	3904
1419	1403	3943	3973	2990	2974
6733	6811	1335	1324	1755	1786
11361	11374	1360	1362	4506	4301
3923	3919	125	<800	3149	3181
3728	3774	1628	1630	95	<800
2678	2713	7946	7991	1863	1786
353	<800	5936	5970	1587	1553
2270	2261	4659	4641	778	774
39391	39409	418	<800	5306	5230
5712	5665	8000	7991	449	<800
637	<800	140	<800	4883	4913
4977	4916	452	<800	1088	1090
6652	6649	1346	1362	6421	6427
7	<800	852	860	5001	5230
3297	3323	890	925		
1835	1818	3158	3181		
10421	10389	6770	6783		
2411	2436	4863	4913		
184	<800	3423	3424		
28	<800	2929	2974		
115	<800	1243	1220		
188	<800	2285	2294		
351	<800	473	<800		
508	<800	964	925		
2446	2436	8680	8644		
1718	1713	1439	1429		
1819	1818	3921	3904		
1729	1818	5549	5552		
413	<800	737	774		
7684	7991	249	<800		

Alignment Scores:  
 Pred. No.: 1,09e-86  
 Score: 90.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 4.75%  
 DB: 9  
 Length: 198470  
 Matches: 90  
 Conservative: 0  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-09-964-956-13 (1-1896) x AC018643 (1-198470)

QY 1340 ValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIle 1359  
|||||  
Db 69968 GTCCCGGCTACCGGACAGGCGTGTGGAGAAAGCCCTGAAGCTCTGGCCAGCTCATC 69909  
QY 1360 AsnAsnLysValPheLeuLeuSerPheLeuArgThrLeuGluSerGlnArgSerPheSer 1379  
|||||  
Db 69908 AACAAACAAGGTGTTCCCTGCTCTCTTCATCGGACGCTTGAGTCGGAGCGTAGCTTCTCC 69849  
QY 1380 MetArgAspArgGlyAsnValAsnLeuIleMetThrValLeuGlnSerLysLeuGlu 1399  
|||||  
Db 69848 ATGCGGCACCGTGGCAACGTCCTCACTCATGACCGGCTGCAGACCAAGCTGGAG 69789  
QY 1400 TyrAlaThrAspValLeuLysGlnLeuAlaAspLeuIleAspLysAsnLeuGluSer 1419  
|||||  
Db 69788 TACGCCAAGTGTGCTGAAGCAGCTGCTGCGCCGAGCTCATTGACAAAGAACCTGGAGAGC 69729  
QY 1420 LysAsnHisProLysLeuLeuLeuArgArg 1429  
|||||  
Db 69728 AAGAACCAAGCTAAGCTGCTCTCAGGAGG 69699

Search completed: June 28, 2003, 23:01:01  
Job time : 12647 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 15:58:05 ; Search time 831 Seconds  
(without alignments)  
5138.133 Million cell updates/sec

Title: US-09-964-956-13  
Perfect score: 1896  
Sequence: 1 MKAMPWNWTCCLSHLLMVG.....QKLAYKLEQVITLMSLDSNK 1896

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 436877

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09964956/runat\_24062003\_124827\_7346/app\_query.fasta\_1.2055  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=oli.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09964956 -CGN\_1\_1\_578 -runat\_24062003\_124827\_7346 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	254	13.9	2597	24	CDNA encoding huma
C	173	9.1	598	22	Human brain expres
C	173	9.1	598	22	Human bone marrow
C	173	9.1	598	22	Probe #11431 used
C	160	8.4	495	22	Human brain expres
C	160	8.4	495	22	Human bone marrow
C	160	8.4	495	22	Probe #24492 used
C	160	8.4	495	22	Human cDNA encodin
8	124	6.5	601	22	Human NOV/ploxin-A
9	63	3.3	1088	22	DNA encoding novel
10	61	3.2	477	23	DNA encoding novel
11	61	3.2	1314	23	Mouse ischaemic co
12	51	2.7	6730	24	Drosophila melanog
13	49	2.6	6382	23	Human cDNA clone (
14	44	2.3	768	22	Human cDNA sequenc
15	44	2.3	1767	22	Human cDNA coding
16	44	2.3	6387	24	Human secreted exp
17	38	2.0	418	21	Human prostate exp
18	33	1.7	995	23	Human prostate exp
19	33	1.7	995	23	Human secreted pro
20	24	1.3	570	21	Bovine embryonic g
21	23	1.2	932	24	Human brain expres
22	18	0.9	306	22	Human genome-deriv
23	18	0.9	306	22	Human genome-deriv
24	18	0.9	306	24	Human genome-deriv
25	18	0.9	518	22	Human brain expres
26	18	0.9	518	22	Probe #1108 used
27	18	0.9	518	24	Human genome-deriv
28	16	0.8	201	22	Human genome-deriv
29	16	0.8	201	24	Human genome-deriv
30	16	0.8	600	22	Human genome-deriv
31	16	0.8	600	24	Human genome-deriv
32	13	0.7	2414	22	Human full-length
33	12	0.6	569	24	Human ovarian canc
34	12	0.6	1844	24	Human ovarian anti
35	12	0.6	2477	20	Protein PRQ235 CDN
36	12	0.6	2477	21	Human PRO235 prote
37	12	0.6	2477	21	Human PRO235 CDNA
38	12	0.6	2477	22	Human cDNA sequenc
39	12	0.6	2477	22	Human PRO235 CDNA
40	12	0.6	6408	22	Human cDNA encodin
41	11	0.6	154	23	Human prostate exp
42	11	0.6	205	23	Human prostate exp
43	11	0.6	205	23	Human prostate exp
44	11	0.6	212	22	Human brain expres
45	11	0.6	212	24	Human genome-deriv

ALIGNMENTS

RESULT 1  
ID ABK70006  
ID ABK70006 standard; DNA; 2597 BP.  
XX  
AC ABK70006;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE cDNA encoding human Pro peptide #46.  
XX  
KW Human; ss; gene; PRO; secreted protein; transmembrane protein;  
XX genetic disorder; tumour; cancer.  
OS Homo sapiens.  
PN WO200224888-A2.  
XX  
PD 28-MAR-2002.



```
PN WO200157275-A2.
XX
PD
PE
PF
PG
PH
PI
PJ
PK
PL
PM
PN
PO
PP
PQ
PR
PS
PT
PU
PV
PW
PX
PY
PZ
QA
QB
QC
QD
QE
QF
QG
QH
QI
QJ
QK
QL
QM
QN
QO
QP
QQ
QR
QS
QT
QU
QV
QW
QX
QY
QZ
RA
RB
RC
RD
RE
RF
RG
RH
RI
RJ
RK
RL
RM
RN
RO
RP
RQ
RR
RS
RT
RU
RV
RW
RX
RY
RZ
SA
SB
SC
SD
SE
SF
SG
SH
SI
SJ
SK
SL
SM
SN
SO
SQ
SR
SS
ST
SU
SV
SW
SX
SY
SZ
TA
TB
TC
TD
TE
TF
TG
TH
TI
TJ
TK
TL
TM
TN
TO
TP
TQ
TR
TS
TT
TU
TV
TW
TX
TY
TZ
UA
UB
UC
UD
UE
UF
UG
UH
UI
UJ
UK
UL
UM
UN
UO
UP
UQ
UR
US
UT
UU
UV
UW
UX
UY
UZ
VA
VB
VC
VD
VE
VF
VG
VH
VI
VJ
VK
VL
VM
VN
VO
VP
VQ
VR
VS
VT
VU
VV
VW
VX
VY
VZ
WA
WB
WC
WD
WE
WF
WG
WH
WI
WJ
WK
WL
WM
WN
WO
WP
WQ
WR
WS
WT
WU
WV
WW
WX
WY
WZ
XA
XB
XC
XD
XE
XF
XG
XH
XI
XJ
XK
XL
XM
XN
XO
XP
XQ
XR
XS
XT
XU
XV
XW
XX
XY
XZ
YA
YB
YC
YD
YE
YF
YG
YH
YI
YJ
YK
YL
YM
YN
YO
YP
YQ
YR
YS
YT
YU
YV
YW
YX
YY
YZ
ZA
ZB
ZC
ZD
ZE
ZF
ZG
ZH
ZI
ZJ
ZK
ZL
ZM
ZN
ZO
ZP
ZQ
ZR
ZS
ZT
ZU
ZV
ZW
ZX
ZY
ZZ
```

141 LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu 160  
|||||  
99 CTGAGGCTGGAGGACCTCTTCAAGCTGGGGAGCCTTATCATAGAAGGAGCACTATCTG 40  
161 SerGlyValAsnGluSerGlySerValPheGlyValle 173  
|||||  
39 TCAGGTGTCAACGAGAGCGGCTCGGTCTTTGGAGTGATC 1  
|||||  
RESULT 3  
AAK36939/C  
ID AAK36939 standard; DNA; 598 BP.  
XX  
AC AAK36939;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed single exon probe SEQ ID NO: 11496.  
XX  
KW Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00668.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
(MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PI WPI; 2001-488900/53.  
XX  
DR Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human bone marrow -  
XX  
PS Example 4; SEQ ID NO: 11496; 658pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
CC the probes of the invention.  
XX  
SQ Sequence 598 BP; 121 A; 169 C; 179 G; 129 T; 0 other;  
Alignment Scores:  
Pred. No.: 9,85e-162 Length: 598  
Score: 173.00 Matches: 173  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.12% Indels: 0  
DB: 22 Gaps: 0  
US-09-964-956-13 (1-1896) x AAK11122 (1-598)  
QY 1 MetLysAlaMetProTyrAsnTrpThrCysLeuSerHisLeuMetValGlyMet 20  
|||||  
519 ATGAAGCCATGCGCTGGAACTGGACTGCTTCTCTCCACCTCTCTCATGGTGGGCGATG 460  
21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
|||||  
459 GGCTCTCCACTTTGTCTACCCGGCAGCCAGCCGCTGTCCAGAGAGAGCGGCTCATTT 400  
41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
|||||  
399 GTACATTCGAGAGAGAGCCGCGAGGTTTCAATCACCTGTGTGGATGAGAGACA 340  
61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
|||||  
339 GGACACATTTACTTGGGGGCGCGTCAATCGGATTTTACANGCTCTCCAGCGACCTGAAGGTC 280  
81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100  
|||||  
279 TTGGTGACGATCAGACAGCGCGGAGGAGACACCCCAAGTGTATCCACCCCGCGATC 220  
101 ValGlnThrCysAsnGluProLeuThrThrThrAsnValAsnLysMetLeuLeuLe 120  
|||||  
219 GTCCAGACCTGCAATGAGCCCTTGACACACCAACAATGTCAACAAGATGCTCCTCATA 160  
121 AspTyrLysGluAsnArgLeuLeuAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
|||||  
159 GACTACAAGAGAACAAGGCTGATTGCTGTGGGAGGCTGTACCAAGGCGATCTGCAAGCTG 100

Human bone marrow expressed single exon probe SEQ ID NO: 11496.  
Human; bone marrow expressed exon; gene expression analysis; probe;  
microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
Homo sapiens.  
WO200157276-A2.  
09-AUG-2001.  
30-JAN-2001; 2001WO-US00668.  
04-FEB-2000; 2000US-0180312.  
26-MAY-2000; 2000US-0207456.  
30-JUN-2000; 2000US-0608408.  
03-AUG-2000; 2000US-0632366.  
21-SEP-2000; 2000US-0234687.  
27-SEP-2000; 2000US-0236359.  
04-OCT-2000; 2000GB-0024263.  
(MOLE-) MOLECULAR DYNAMICS INC.  
Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2001-483446/52.  
Single exon nucleic acid probes for analyzing gene expression in human  
brains -  
Example 4; SEQ ID NO: 11113; 650pp + Sequence Listing; English.  
The present invention provides a number of single exon nucleic acid  
probes which are derived from genomic sequences expressed in the human  
brain. They can be used to measure gene expression in brain cell samples,  
which may enable the diagnosis and improved treatment of nervous system  
diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
epilepsy and cancers. The present sequence is one of the probes of the  
invention.  
Sequence 598 BP; 121 A; 169 C; 179 G; 129 T; 0 other;  
Alignment Scores:  
Pred. No.: 9,85e-162 Length: 598  
Score: 173.00 Matches: 173  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.12% Indels: 0  
DB: 22 Gaps: 0  
US-09-964-956-13 (1-1896) x AAK11122 (1-598)  
QY 1 MetLysAlaMetProTyrAsnTrpThrCysLeuSerHisLeuMetValGlyMet 20  
|||||  
519 ATGAAGCCATGCGCTGGAACTGGACTGCTTCTCTCCACCTCTCTCATGGTGGGCGATG 460  
21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
|||||  
459 GGCTCTCCACTTTGTCTACCCGGCAGCCAGCCGCTGTCCAGAGAGAGCGGCTCATTT 400  
41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
|||||  
399 GTACATTCGAGAGAGAGCCGCGAGGTTTCAATCACCTGTGTGGATGAGAGACA 340  
61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
|||||  
339 GGACACATTTACTTGGGGGCGCGTCAATCGGATTTTACANGCTCTCCAGCGACCTGAAGGTC 280  
81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100  
|||||  
279 TTGGTGACGATCAGACAGCGCGGAGGAGACACCCCAAGTGTATCCACCCCGCGATC 220  
101 ValGlnThrCysAsnGluProLeuThrThrThrAsnValAsnLysMetLeuLeuLe 120  
|||||  
219 GTCCAGACCTGCAATGAGCCCTTGACACACCAACAATGTCAACAAGATGCTCCTCATA 160  
121 AspTyrLysGluAsnArgLeuLeuAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
|||||  
159 GACTACAAGAGAACAAGGCTGATTGCTGTGGGAGGCTGTACCAAGGCGATCTGCAAGCTG 100

Db 459 GGCTCTCCACTTTGGCTCACCAGGAGCCAGCCCGCTGTCTCCAGAAACAGCGGTCAATTT 400  
Qy 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
Db 399 GTCACATTCGAGGAGAGCGCGCGAGGGTTTCAATCACCTGGTGGTGATGAGAGGACA 340  
Qy 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
Db 339 GGACACATTTACTTGGGGCGCGTCAATCGGATTTACAAAGCTCTCCAGCGACTGAAGGTC 280  
Qy 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100  
Db 279 TTGTGTGCGCATGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 220  
Qy 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIle 120  
Db 219 GTCCAGACCTGCATGAGCCCTCACCACACCAACATGTCACAAAGATGCTCTCTCATA 160  
Qy 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
Db 159 GACTACAGGAGAGACAGGCTGATTGCGCTGTGGAGCCTGTACCAAGGCATCTGCAAGCTG 100  
Qy 141 LeuArgLeuGluAspPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu 160  
Db 99 CTGAGGCTGGAGGACCTCTCAAGCTGGGGAGCCTTATCATAGAAGAGGACACTATCTG 40  
Qy 161 SerGlyValAsnGluSerGlySerValPheGlyValIle 173  
Db 39 TCAGGTCTCAACGAGAGCGGCTCGGTCTTTGGAGTGATC 1

## RESULT 4

AAI42745/c

ID AAI42745 standard; DNA; 598 BP.

XX AC AAI42745;

XX AC AAI42745;

XX DT 17-OCT-2001 (first entry)

XX XX Probe #11431 used to measure gene expression in human placenta sample.  
XX DE Probe; microarray; human; placenta; antenatal diagnosis;  
XX KW genetic disorder; ss.

XX KW Homo sapiens.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX XX (MOLE-) MOLECULAR DYNAMICS INC.

XX XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX WPI; 2001-488897/53.

XX XX Human genome-derived single exon nucleic acid probes useful for

XX XX analyzing gene expression in human placenta

XX XX Claim 25; SEQ ID No 11431; 654pp; English.

XX XX The present invention relates to single exon nucleic acid probes (SENP).

XX CC The present sequence is one such probe. The probes are useful for

XX CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful  
CC for antenatal diagnosis of human genetic disorders.

SQ Sequence 598 BP; 121 A; 169 C; 179 G; 129 T; 0 other;

## Alignment Scores:

Pred. No.: 9.85e-162 Length: 598  
Score: 173.00 Matches: 173  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.12% Indels: 0  
DB: 22 Gaps: 0

US-09-964-956-13 (1-1896) x AAI42745 (1-598)

Qy 1 MetLysAlaMetProTrpAsnTrpThrCysLeuSerHisLeuMetValGlyMet 20  
Db 519 ATGAAGCCATGCCCTGGAACTGGACCTGCTCTCTCCACCTCCCTCATGGTGGGATG 460  
Qy 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
Db 459 GGCTCTCCACTTTGGCTCACCAGGAGCCAGCCCGCTGTCCAGAAAGCAGCGGTCAATTT 400  
Qy 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
Db 399 GTCACATTCGAGGAGAGCGCGGAGGGTTTCAATCACCTGGTGGTGATGAGAGGACA 340  
Qy 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
Db 339 GGACACATTTACTTGGGGCGCGTCAATCGGATTTACAAAGCTCTCCAGGACCTGAAGGTC 280  
Qy 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100  
Db 279 TTGTGTGCGCATGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 220  
Qy 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIle 120  
Db 219 GTCCAGACCTGCATGAGCCCTCACCACACCAACATGTCACAAAGATGCTCTCTCATA 160  
Qy 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
Db 159 GACTACAGGAGAGACAGGCTGATTGCGCTGTGGAGCCTGTACCAAGGCATCTGCAAGCTG 100  
Qy 141 LeuArgLeuGluAspPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu 160  
Db 99 CTGAGGCTGGAGGACCTCTTCAAGCTGGGGAGCCTTATCATAGAAGAGGAGCAGCTATCTG 40  
Qy 161 SerGlyValAsnGluSerGlySerValPheGlyValIle 173  
Db 39 TCAGGTCTCAACGAGAGCGGCTCGGTCTTTGGAGTGATC 1

## RESULT 5

AAK23792/c

ID AAK23792 standard; DNA; 495 BP.

XX AC AAK23792;

XX AC AAK23792;

XX DT 05-NOV-2001 (first entry)

XX DE Human brain expressed single exon probe SEQ ID NO: 23783.

XX KW Human; brain expressed exon; gene expression analysis; probe;

XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX KW epilepsy; cancer; ss.

XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX XX

```
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 23783; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 495 BP; 96 A; 132 C; 153 G; 114 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 6.7e-149 Length: 495
XX Score: 160.00 Matches: 160
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 8.44% Indels: 0
XX DB: 22 Gaps: 0
XX
US-09-964-956-13 (1-1896) x AAK23792 (1-495)
QY 1 MetLysAlaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMet 20
Db 481 ATGAAGCCATGCCTGGAACTGGACCTGCTTCTCCACCTCCCTCATGGTGGGCATG 422
QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40
Db 421 GGCTCCTCCACTTTGTCTACCCGGCAGCCAGCCCGCTGTCCAGGACCTCCAGGCTCATTT 362
QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60
Db 361 GTCACATTCGAGGAGAGCCCGGAGGGTTTCAATCACCTGGTGGTGGATGAGAGGACA 302
QY 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80
Db 301 GGACACATTTACTTGGGGCGGTCAATCGGATTTACAGCTTCCAGGACCTCGAGGTC 242
QY 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100
Db 241 TTGGTGACGATGAGACAGCGCGGAGGACAAACCCCAAGTGTACCCACCCCGCATC 182
QY 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIle 120
Db 181 GTCAGACCTGCAATGAGCCCTTGACCACCAACCAATGTCAACAAGATGCTCCTCATA 122
QY 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140
Db 121 GACTACAGGAGAACAGGCTGATGGCTGTGGAGCCTGTACCAAGCATCTGCAAGCTG 62
QY 141 LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu 160
Db 61 CTGAGGCTGGAGGACCTCTTCAAGCTGGGGAGCCTTATCATATAAGAGGACCACTATCTG 2
XX
XX RESULT 6
XX AAK49882/c
```

```
ID AAK49882 standard; DNA; 495 BP.
XX
XX AAK49882;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 24439.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 24439; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX Sequence 495 BP; 96 A; 132 C; 153 G; 114 T; 0 other;
XX
```

```
Alignment Scores:
Pred. No.: 6.7e-149 Length: 495
Score: 160.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.44% Indels: 0
DB: 22 Gaps: 0
```

US-09-964-956-13 (1-1896) x AAK49882 (1-495)

```
QY 1 MetLysAlaMetProTrpAsnTrpThrCysLeuSerHisLeuLeuMetValGlyMet 20
Db 481 ATGAAGCCATGCCTGGAACTGGACCTGCTTCTCCACCTCCCTCATGGTGGGCATG 422
QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40
Db 421 GGCTCCTCCACTTTGTCTACCCGGCAGCCAGCCCGCTGTCCAGGACCTCCAGGCTCATTT 362
QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60
Db 361 GTCACATTCGAGGAGAGCCCGGAGGGTTTCAATCACCTGGTGGTGGATGAGAGGACA 302
QY 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80
Db 301 GGACACATTTACTTGGGGCGGTCAATCGGATTTACAGCTTCCAGGACCTCGAGGTC 242
```

Qy 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100  
 Db 241 TTGTGACGATGAGACAGGCGCGAGGACAAACCCAAAGTGTACCCACCCGCGATC 182  
 Qy 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeu 120  
 Db 181 GTCCAGACCTGCAATGAGCCCTGACCACCAACCAATGTCAACAAGATGCTCCTCATA 122  
 Qy 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
 Db 121 GACTACAGGAGACAGCGCTGATTCCTGTGGAGCCTGTACCAAGGATCTGCAAGCTG 62  
 Qy 141 LeuArgLeuGluAspLeuLysLeuGluGluProTyrHisLysLysLysGluHisTyrLeu 160  
 Db 61 CTGAGGCTGGAGGACCTCTCAAGCTGGGGAGCCTTATCATAGAAGGAGCACTATCTG 2

## RESULT 7

AAI55806/C

ID AAI55806 standard; DNA; 495 BP.

AC AAI55806;

XX 17-OCT-2001 (first entry)

XX Probe #24492 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder; ss.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -

XX Claim 25; SEQ ID NO 24492; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.

XX Sequence 495 BP; 96 A; 132 C; 153 G; 114 T; 0 other;

## Alignment Scores:

Pred No.: 6'7e-149 Length: 495  
 Score: 160.00 Matches: 160  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 8.44% Indels: 0  
 DB: 22 Gaps: 0

US-09-964-956-13 (1-1896) x AAI55806 (1-495)

Qy 1 MetLysAlaMetProTirPAsnTirPThrCysLeuSerHisLeuLeuMetValGlyMet 20  
 Db 481 ATGAAGCCATGCCCTGGAACTGGACCTGCCTTCTCTCCACCTCTCATGGTGGGCATG 422  
 Qy 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
 Db 421 GGCTCCTCCACTTTGCTCACCGGCGAGCGCCGCTGTCCAGAACGACGCGTCAATTT 362  
 Qy 41 ValThrPheArgGlyCyluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
 Db 361 GTCACATTCGAGAGGAGCGCGCGAGGGTTTCAATCACCTGGTGGTGGATGAGAGACA 302  
 Qy 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
 Db 301 GCACACATTTACTTGGGGCGCGTCANTCGATTACAGCTCTCCAGCGACCTGAAGTCT 242  
 Qy 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100  
 Db 241 TTGGTGACGATGAGACAGGCGCGAGGACAAACCCCAAGTGTACCCACCCGCGATC 182  
 Qy 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeu 120  
 Db 181 GTCCAGACCTGCAATGAGCCCTGACCACCAACAATGTCAACAAGATGCTCCTCATA 122  
 Qy 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
 Db 121 GACTACAAGGAGAACAGGCTGATTGCTGTGGAGCCTGTACCAAGGATCTGCAAGCTG 62  
 Qy 141 LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysLysGluHisTyrLeu 160  
 Db 61 CTGAGGCTGGAGGACCTCTTCAAGCTGGGGAGCCTTATCATAGAAGGAGCACTATCTG 2

## RESULT 8

AAS00023

ID AAS00023 standard; cDNA; 601 BP.

XX AAS00023;

XX 09-MAY-2001 (first entry)

XX Human cDNA encoding Plexin A-4.

XX Human; Plexin A-4; semaphorin domain; hyperplasia; neoplasia; cancer;  
 KW neurodegenerative disease; autoimmune disease; lupus; multiple sclerosis;  
 KW inflammatory bowel disease; diabetes type I; rheumatoid arthritis;  
 KW immunogen; antibody; ss.

XX .Homo sapiens.

XX Key Location/Qualifiers

FH 2..601

CDS /\*tag= a

FT /transl\_except= (pos:539..544,aa:Thr)

FT /partial

FT /product= "Plexin A-4"

FT /note= "No start or stop codon"

XX WO200114420-A2.

XX 01-MAR-2001.

XX 25-AUG-2000; 2000WO-US23365.

XX 25-AUG-1999; 99US-0150576.

XX (UYTO-) UNIV TORINO.

XX (REGC) UNIV CALIFORNIA.

XX Artigiani S, Comoglio PM, Goodman CS, Tesier-Lavigne M;  
 PI Tamagnone L;

XX WPI; 2001-226610/23.



CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention.

XX Sequence 1088 BP; 256 A; 339 C; 310 G; 183 T; 0 other;

Alignment Scores:  
Pred. No.: 2.98e-52 Length: 1088  
Score: 63.00 Matches: 63  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.32% Indels: 0  
DB: 22 Gaps: 0

US-09-964-956-13 (1-1896) x ABA08997 (1-1088)

QY 1665 GluclyAspArgGlySerLysMetValSerGluIleThrLeuThrArgLeuLeuAlaThr 1684  
DB 759 GAGGGTGACCGCGCAGCAAGATGGTCTCGGAGATCTACTTGACACAGCGCTACTGGCCACC 818  
QY 1685 LysGlyThrLeuGlnLysPheValAspLeuPheGluThrIlePheSerThrAlaHis 1704  
DB 819 AAGGCGACACTGCAGAGTTTGTGGACGACCTGTTGAGACCATCTTCAGCACGSCACAC 878  
QY 1705 ArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAla 1724  
DB 879 CGGGGCTCAGCCCTGCCGCTGCCATCAAGTACATGTTTCGATCTCCTCGATGAGCAGGCC 938  
QY 1725 AsPlyHis 1727  
DB 939 GACAAGCAC 947

RESULT 10

AAS68294

ID AAS68294 standard; cDNA; 477 BP.

XX AAS68294;

AC AAS68294;

XX 13-FEB-2002 (first entry)

DT DNA encoding novel human diagnostic protein #4098.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG04107.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

PS Claim 1; SEQ ID No 4098; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy involving  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 477 BP; 108 A; 123 C; 128 G; 118 T; 0 other;

Alignment Scores:  
Pred. No.: 1.33e-50 Length: 477  
Score: 61.00 Matches: 61  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.22% Indels: 0  
DB: 23 Gaps: 0

US-09-964-956-13 (1-1896) x AAS68294 (1-477)

QY 397 LeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyValSer 416

DB 118 CTCCTTAACCATTCACGATTAACCTCTGTGGCTGGACATGAATGCTCCCTGGAGTGCTCC 177

QY 417 AspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThrSerValIle 436

DB 178 GACATGGTGGTGAATTCCTCGTCTTACGGAGGACAGGACCGCATGAGCTGTGTCATC 237

QY 437 AlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLys 456

DB 238 GCATATGTCTACAAGAACCACTCTCTGGCCCTTTGTGGGCACCAAAAGTGCAAGCTCAAG 297

QY 457 Lys 457

DB 298 AAG 300

RESULT 11

AAS68296

ID AAS68296 standard; cDNA; 1314 BP.

XX AAS68296;

AC AAS68296;

XX 13-FEB-2002 (first entry)

DT DNA encoding novel human diagnostic protein #4100.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS



```
Db 5962 CCCAGTTATAGAACTGGGTAGAAAGATCTAT 5994
|||||
RESULT 13
ABLI17265
ID ABLI17265 standard; DNA; 6382 BP.
XX
AC ABLI17265;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 3268.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 3268; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
XX sequences (ABLI01840-ABLI16175) and the encoded proteins
XX (AB57737-AB572072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6382 BP; 2073 A; 1195 C; 1282 G; 1832 T; 0 other;

Alignment Scores:
Pred. No.: 1,266-37 Length: 6382
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2,58% Indels: 0
DB: 23 Gaps: 0

US-09-964-956-13 (1-1896) x ABLI17265 (1-6382)

QY 1673 ValSerGluIleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheVal 1692
|||||
Db 5311 GTGTCAGAGATTACTTACCAGGCTATTGGCTACTAAAGAACGCTTCAGAAATTGTA 5370
|||||

QY 1693 AspAspLeuPheGluThrIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAla 1712
|||||
Db 5371 GATGATCTCTTTGAACAGCATATTAGCACCGCTCACCGTGGATCGGCACTTCCATTAGCT 5430
|||||

QY 1713 IleLysTyrMetPheAspPheLeuAsp 1721
|||||
```

```
Db 5431 ATAAAGTACATGTTTCATTTTCTTGAC 5457
RESULT 14
AAH03972
ID AAH03972 standard; cDNA; 768 BP.
XX
AC AAH03972;
XX
XX 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:807.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
KW Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 1; SEQ ID 807; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesising 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primers are useful for synthesising polynucleotides,
XX in gene therapy. The primers are also useful for the
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
SQ Sequence 768 BP; 175 A; 222 C; 207 G; 161 T; 3 other;

Alignment Scores:
Pred. No.: 1,596-33 Length: 768
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2,32% Indels: 0
```

DB: 22 Gaps: 0

US-09-964-956-13 (1-1896) x AAH03972 (1-768)

QY 1765 SerIleThrAspAlaCysLeuSerValValAlaGlnThrPheMetAspSerCysSerThr 1784  
|||||  
Db 234 AGCATCACGGACGCTCTCTGTGGTGGCCAGACCTTCATGGACTCTTTGTTCAAG 293  
|||||

QY 1785 SerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysAspIle 1804  
|||||  
Db 294 TCAGAGCACCGCTGGCGAAGAGTCCCTCCACACAGCTGCTATATGCCAAGGACATC 353  
|||||

QY 1805 ProSerTyrLys 1808  
|||||  
Db 354 CCCAGCTACAAG 365  
|||||

RESULT 15

AAH15398

ID AAH15398 standard; cDNA: 1767 BP.

AC AAH15398;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13599.

XX

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

XX

XX 28-JUL-2000; 2000EP-0116126.

XX

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX

XX (HELI-) HELIX RES INST.

XX

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

XX WPI; 2001-318749/34.

Search completed: June 28, 2003, 19:31:03  
Job time : 896 secs

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
XX of the present invention.

SQ Sequence 1767 BP; 461 A; 446 C; 431 G; 429 T; 0 other;

Alignment Scores:

Pred. No.: 3 47e-33 Length: 1767  
Score: 44.00 Matches: 44  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.32% Indels: 0  
DB: 22 Gaps: 0

US-09-964-956-13 (1-1896) x AAH15398 (1-1767)

QY 1765 SerIleThrAspAlaCysLeuSerValValAlaGlnThrPheMetAspSerCysSerThr 1784  
|||||

Db 234 AGCATCACGGACGCTCTCTGTGGTGGCCAGACCTTCATGGACTCTTTGTTCAAG 293  
|||||

QY 1785 SerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysAspIle 1804  
|||||

Db 294 TCAGAGCACCGCTGGCGAAGAGTCCCTCCACACAGCTGCTATATGCCAAGGACATC 353  
|||||

QY 1805 ProSerTyrLys 1808  
|||||

Db 354 CCCAGCTACAAG 365  
|||||

Claim 8; SEQ ID 13599; 2537pp + CD ROM; English.

XX

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesising polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 19:19:01 ; Search time 173 Seconds  
(without alignments)  
3361.035 Million cell updates/sec

Title: US-09-964-956-13

Perfect score: 1896  
Sequence: 1 MKAMPNWTCLLSHLLVMGM.....OKLAYKLEQVITLMSLDSNK 1896

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 878600

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09964956/runat\_24062003\_124829\_7384/app\_query.fasta\_1.2055  
-DB=Issued Patents NA -OEMT=fastap -SUFFIX=oli.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdd  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -NAXLEN=2000000000  
-USER=US09964956.ecgn\_1\_1\_78\_erunat\_24062003\_124829\_7384 -NCPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAEXT=60 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=60 -YGAEXT=60 -DELOP=6 -DEEXT=7

Database :

Issued Patents\_NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	0.5	4707	3	US-09-181-706-1
2	10	0.5	4707	4	US-09-458-791-1
3	10	0.5	4707	4	US-09-459-066-1
4	8	0.4	1220	4	US-09-227-357-54
5	8	0.4	1353	4	US-09-028-274A-17
6	8	0.4	1631	1	US-08-663-310-12
7	8	0.4	1631	2	US-09-006-491-12
8	8	0.4	1631	3	US-09-335-919-12
9	8	0.4	1692	4	US-08-906-743-1
10	8	0.4	2071	4	US-08-816-088-1
11	8	0.4	2176	4	US-09-399-913-39
12	8	0.4	2176	4	US-09-298-731-39

13	8	0.4	2219	1	US-08-606-322-1	Sequence 1, Appli
14	8	0.4	2841	4	US-09-399-913-31	Sequence 31, Appl
15	8	0.4	4605	4	US-09-221-017B-128	Sequence 128, App
16	8	0.4	9551	1	US-08-056-200-93	Sequence 93, Appl
17	8	0.4	9551	2	US-08-800-644-93	Sequence 93, Appl
18	8	0.4	12720	1	US-08-403-866-11	Sequence 11, Appl
19	7	0.4	28	1	US-08-627-845-10	Sequence 10, Appl
20	7	0.4	28	1	US-08-627-845-10	Sequence 10, Appl
21	7	0.4	30	1	US-08-465-798-10	Sequence 10, Appl
22	7	0.4	30	4	US-09-271-437-5	Sequence 5, Appli
23	7	0.4	33	2	US-08-177-109A-3	Sequence 3, Appli
24	7	0.4	33	2	US-08-687-706-3	Sequence 3, Appli
25	7	0.4	34	1	US-08-644-664B-36	Sequence 36, Appl
26	7	0.4	34	2	US-08-761-277A-36	Sequence 36, Appl
27	7	0.4	36	3	US-09-361-434-30	Sequence 30, Appl
28	7	0.4	36	4	US-09-635-025-30	Sequence 30, Appl
29	7	0.4	39	2	US-08-684-594-14	Sequence 14, Appl
30	7	0.4	50	4	US-09-313-221A-84	Sequence 84, Appl
31	7	0.4	69	4	US-09-020-846-56	Sequence 56, Appl
32	7	0.4	146	4	US-09-585-651A-11	Sequence 11, Appl
33	7	0.4	146	4	US-09-585-651A-11	Sequence 12, Appl
34	7	0.4	162	1	US-08-053-867A-4	Sequence 4, Appli
35	7	0.4	166	1	US-07-973-321B-2	Sequence 2, Appli
36	7	0.4	166	1	US-08-090-527A-2	Sequence 2, Appli
37	7	0.4	209	1	US-08-510-032A-8	Sequence 8, Appli
38	7	0.4	209	3	US-08-688-514-8	Sequence 8, Appli
39	7	0.4	219	4	US-08-729-601A-47	Sequence 47, Appl
40	7	0.4	220	3	US-08-847-844A-91	Sequence 91, Appl
41	7	0.4	231	4	US-09-605-785-465	Sequence 465, App
42	7	0.4	231	4	US-09-439-313-465	Sequence 465, App
43	7	0.4	231	4	US-09-352-616A-465	Sequence 465, App
44	7	0.4	261	4	US-09-605-785-411	Sequence 411, App
45	7	0.4	261	4	US-09-439-313-411	Sequence 411, App

ALIGNMENTS

RESULT 1

US-09-181-706-1  
; Sequence 1, Application US/09181706  
; Patent No. 6130068  
; GENERAL INFORMATION:  
; APPLICANT: Melanie K. Spriggs, Michael R. Comeau,  
; APPLICANT: Robert F. DuBoise, Richard S. Johnson  
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Janis C. Henry  
; STREET: 51 University St.  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/181,706  
; FILING DATE: October 28, 1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/958,598 (converted to a  
; APPLICATION NUMBER: Provisional, see below)  
; FILING DATE: October 28, 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: --to be assigned-- (USN 08/958,598  
; APPLICATION NUMBER: conversion to Provisional application)  
; FILING DATE: October 26, 1998  
; CLASSIFICATION:

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4707 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4707
; US-09-181-706-1
;
; Alignment Scores:
; Pred. No.: 7.77 Length: 4707
; Score: 10.00 Matches: 10
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 0.53% Indels: 0
; DB: Gaps: 0
;
; US-09-964-956-13 (1-1896) x US-09-181-706-1 (1-4707)
;
; Qy 1796 AsnLysLeuLeuTyAlaLysAspIlePro 1805
; Db 4399 AATAAGCTTCTCTATGCCAAGGATATCCCA 4428
;
; RESULT 2
; US-09-458-791-1
; Sequence 1, Application US/09458791
; Patent No. 6174689
; GENERAL INFORMATION:
; APPLICANT: Spriggs, Melanie
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; RECEPTOR DNA AND POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS/Windows 95
; SOFTWARE: Word for Windows 95, 7.0a
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/458,791
; FILING DATE: 10-Dec-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,598
; FILING DATE: 28-OCT-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4707 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4707
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-09-458-791-1
;
; Alignment Scores:
; Pred. No.: 7.77 Length: 4707
; Score: 10.00 Matches: 10
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 0.53% Indels: 0
; DB: Gaps: 0
;
; US-09-964-956-13 (1-1896) x US-09-458-791-1 (1-4707)
;
; Qy 1796 AsnLysLeuLeuTyAlaLysAspIlePro 1805
; Db 4399 AATAAGCTTCTCTATGCCAAGGATATCCCA 4428
;
; RESULT 3
; US-09-459-066-1
; Sequence 1, Application US/09459066
; Patent No. 6187909
; GENERAL INFORMATION:
; APPLICANT: Spriggs, Melanie
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; RECEPTOR DNA AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS/Windows 95
; SOFTWARE: Word for Windows 95, 7.0a
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/459,066
; FILING DATE:
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,598
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4707 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4707
```

```

US-09-459-066-1
Alignment Scores:
Pred. No.: 7.77 Length: 4707
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.53% Indels: 0
DB: 4 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-459-066-1 (1-4707)
QY 1796 AsnLysLeuLeuTyrlLysAspIlePro 1805
DB 4399 AATAAGCTTCTCTATGCCAAGGATATCCCA 4428

RESULT 4
US-09-227-357-54
; Sequence 54, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947

```

```

; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 1220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1197)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1208)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1209)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-54

Alignment Scores:
Pred. No.: 223 Length: 1220
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.42% Indels: 0
DB: 4 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-227-357-54 (1-1220)
QY 1257 LeulleAlaTyrlLysArgLysSer 1264
DB 93 CTCATAGCTTATAAAAGAAAGTCC 116

RESULT 5
US-09-028-274A-17
; Sequence 17, Application US/09028274A
; Patent No. 6326465
; GENERAL INFORMATION:
; APPLICANT: Hess, Allan D.
; TITLE OF INVENTION: Specificity of Effector T Lymphocytes
; TITLE OF INVENTION: InAutologous Graft-vs-Host Disease: Role for the MHC Class
; FILE REFERENCE: 01107.73711
; CURRENT APPLICATION NUMBER: US/09/028,274A
; CURRENT FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/039,032
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Rattus rattus

```

US-09-028-274A-17

Alignment Scores: 248 Length: 1353  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.42% Indels: 0  
DB: 4 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-028-274A-17 (1-1353)

Qy 392 ProCysSerAlaLeuLeuThr 399

Db 936 CCTGCTCTCAGCCCTCTAC 959

RESULT 6

US-08-663-310-12/c  
; Sequence 12, Application US/08663310  
; Patent No. 5811273  
; GENERAL INFORMATION:  
; APPLICANT: MISAWA, No. 5811273ihiko  
; APPLICANT: KONDO, Keiji  
; APPLICANT: KAJIWARA, Susumu  
; APPLICANT: YOKOYAMA, Akihiro  
; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF  
; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE  
; TITLE OF INVENTION: XANTHOPHYLLS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/663.310  
; FILING DATE: 23-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/JP94/02220  
; FILING DATE: 26-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-235917  
; FILING DATE: 05-SEP-1994

ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 49441/109  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1631 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

US-08-663-310-12  
Alignment Scores: 298 Length: 1631  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.42% Indels: 0

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.42% Indels: 0  
DB: 1 Gaps: 0

US-09-964-956-13 (1-1896) x US-08-663-310-12 (1-1631)

Qy 1422 HisProLysLeuLeuArgArg 1429

Db 101 CATCGAAACTCTCTCGGTAGG 78

RESULT 7

US-09-006-491-12/c  
; Sequence 12, Application US/09006491  
; Patent No. 5972690  
; GENERAL INFORMATION:  
; APPLICANT: MISAWA, No. 5972690ihiko  
; APPLICANT: KONDO, Keiji  
; APPLICANT: KAJIWARA, Susumu  
; APPLICANT: YOKOYAMA, Akihiro  
; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF  
; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE  
; TITLE OF INVENTION: XANTHOPHYLLS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/006,491  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/663,310  
; FILING DATE: 23-SEP-1996  
; APPLICATION NUMBER: WO PCT/JP94/02220  
; FILING DATE: 26-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-235917  
; FILING DATE: 05-SEP-1994

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-348737  
; FILING DATE: 27-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 49441/109  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1631 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

US-09-006-491-12  
Alignment Scores: 298 Length: 1631  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.42% Indels: 0

DB: 2 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-006-491-12 (1-1631)

QY 1422 HisProLysLeuLeuArgArg 1429  
|||||

Db 101 CATCGAAACTCTGCTGCTAGG 78

RESULT 8

US-09-335-919-12/c  
; Sequence 12, Application US/09335919  
; Patent No. 6150130  
; GENERAL INFORMATION:  
; APPLICANT: MISA, NO. 6150130hiho  
; APPLICANT: KONO, Keiji  
; APPLICANT: KAJIWARA, Susumu  
; APPLICANT: YOKOYAMA, Akihito  
; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF  
; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/335,919  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/663,310  
; FILING DATE: 23-SEP-1996  
; APPLICATION NUMBER: WO PCT/JP94/02220  
; FILING DATE: 26-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-235917  
; FILING DATE: 05-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-348737  
; FILING DATE: 27-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 49441/109  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1631 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

US-09-335-919-12

Alignment Scores:  
Pred. No.: 298 Length: 1631  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.42% Indels: 0  
DB: 3 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-335-919-12 (1-1631)

QY 1422 HisProLysLeuLeuArgArg 1429  
|||||

Db 101 CATCGAAACTCTGCTGCTAGG 78

RESULT 9

US-08-906-743-1/c  
; Sequence 1, Application US/08906743  
; Patent No. 6232097  
; GENERAL INFORMATION:  
; APPLICANT: Lawlor, Elizabeth  
; TITLE OF INVENTION: No. 6232097el args  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/906,743  
; FILING DATE: 06-AUG-1997  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 9619071.5  
; FILING DATE: 12-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gimmi, Edward R.  
; REGISTRATION NUMBER: 38,891  
; REFERENCE/DOCKET NUMBER: P31625  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-4478  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1692 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA

US-08-906-743-1

Alignment Scores:  
Pred. No.: 309 Length: 1692  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.42% Indels: 0  
DB: 4 Gaps: 0

US-09-964-956-13 (1-1896) x US-08-906-743-1 (1-1692)

QY 1261 LysArgLysSerArgGluSerAsp 1268  
|||||

Db 135 AAGAGAAAAAGCAGGGAAGCGAT 112

RESULT 10

US-09-816-088-1  
; Sequence 1, Application US/09816088  
; Patent No. 6326180  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME, NUCLEIC ACID  
; TITLE OF INVENTION: MOLECULES ENCODING HUMAN ENZYME, AND USES THEREOF  
; FILE REFERENCE: CL001181  
; CURRENT APPLICATION NUMBER: US/09/816,088

```

; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2071
; TYPE: DNA
; ORGANISM: Human
US-09-816-088-1
Alignment Scores:
Pred. No.: 377 Length: 2071
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.42% Indels: 0
DB: 4 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-816-088-1 (1-2071)
QY 1556 ArgGlnGlySerGlyAlaArgMet 1563
DB 1161 AGGCAGGCTCTGGGGCCAGAAATG 1184

RESULT 11
US-09-399-913-39
; Sequence 39, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Huai-Ping
; APPLICANT: An, Wenglan
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 09/350,614
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 09/350,874
; EARLIER FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 2176
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(124)
US-09-399-913-39
Alignment Scores:
Pred. No.: 396 Length: 2176
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.42% Indels: 0
DB: 4 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-399-913-39 (1-2176)
QY 1099 ProAlaLeuAlaLeuGlyProAsp 1106
DB 1962 CCGGCCTTGGCCTTGGGGCCAGAC 1985

RESULT 12
US-09-298-731-39
; Sequence 39, Application US/09298731
; Patent No. 636197
; GENERAL INFORMATION:
; APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070
; CURRENT APPLICATION NUMBER: US/09/298,731
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 2176
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(124)
US-09-298-731-39
Alignment Scores:
Pred. No.: 396 Length: 2176
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.42% Indels: 0
DB: 4 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-298-731-39 (1-2176)
QY 1099 ProAlaLeuAlaLeuGlyProAsp 1106
DB 1962 CCGGCCTTGGCCTTGGGGCCAGAC 1985

RESULT 13
US-08-606-322-1
; Sequence 1, Application US/08606322
; Patent No. 5753501
; GENERAL INFORMATION:
; APPLICANT: Crueger, Anneliese; Piepersberg,
; APPLICANT: Wolfgang; Distler, Jurgan; and
; APPLICANT: Straumann, Ansgar
; TITLE OF INVENTION: ACARBOSE BIOSYNTHESIS GENES FROM
; TITLE OF INVENTION: ACTINOPLANES SP., PROCESS FOR THE ISOLATION
; TITLE OF INVENTION: THEREOF AND THE USE THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 mb
; MEDIUM TYPE: storage
; COMPUTER: Bravo 3/25s
; OPERATING SYSTEM: DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,322
; FILING DATE: 23-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 07 214.6
; FILING DATE: 02-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: Bayer 9537-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
```

```

; TELEFAX: (914) 332-1844
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2219 Nucleotides
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Actinoplanes sp. SE 50/110
; US-09-606-322-1
Alignment Scores:
Pred. No.: 404 Length: 2219
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.42% Indels: 0
DB: 1 Gaps: 0

US-09-964-956-13 (1-1896) x US-08-606-322-1 (1-2219)
Oy 1679 ThrArgLeuAlaThrlyGly 1686
Db 1683 ACCAGGTGCTGCCACGAAGGG 1706

RESULT 14
US-09-399-913-31
; Sequence 31, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Hui-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 09/350,614
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 09/350,874
; EARLIER FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 2841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(768)
US-09-399-913-31
Alignment Scores:
Pred. No.: 516 Length: 2841
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.42% Indels: 0
DB: 4 Gaps: 0

US-09-964-956-13 (1-1896) x US-09-399-913-31 (1-2841)
Oy 1099 ProAlaLeuAlaLeuGlyProAsp 1106
Db 2506 CCGGCTTGGCTTGGGCCGAC 2629

RESULT 15
US-09-221-017B-128
; Sequence 128, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4605 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...4605
US-09-221-017B-128
Alignment Scores:
Pred. No.: 832 Length: 4605
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.42% Indels: 0
DB: 4 Gaps: 0
```

US-09-964-956-13 (1-1896) x US-09-221-017B-128 (1-4605)

Qy 1193 ThrValSeraspValGlnLeuLeu 1200

Db 3036 ACTGCTCGGACGTGCAGCTTCTA 3059

Search completed: June 29, 2003, 00:02:38  
Job time : 293 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 23:01:06 ; Search time 563 Seconds  
(without alignments)  
4999.137 Million cell updates/sec

Title: us-09-964-956-13

Perfect score: 1896

Sequence: 1 MKAMPWNWTCLLSHLLMVG.....QKLYAKLEQVITLMSLDSNK 1896

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2108995

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US0964956/runat\_24062003\_124832\_7501/app.query.fasta\_1.2055  
-DB=Published Applications\_NA -QFMT=fastap -SUFFIX=oli.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.coi -LIST=45 -DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US0964956.cgn2\_1\_1\_376\_runat\_24062003\_124832\_7501  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	264	13.9	2597	9	US-10-245-103-91 Sequence 91, Appl
2	264	13.9	2597	9	US-10-245-107-91 Sequence 91, Appl
3	264	13.9	2597	9	US-10-245-143-91 Sequence 91, Appl
4	264	13.9	2597	9	US-10-245-771-91 Sequence 91, Appl

5	264	13.9	2597	9	US-10-245-851-91 Sequence 91, Appl
6	264	13.9	2597	9	US-10-245-883-91 Sequence 91, Appl
7	264	13.9	2597	9	US-10-237-535-91 Sequence 91, Appl
8	264	13.9	2597	9	US-10-238-183-91 Sequence 91, Appl
9	264	13.9	2597	9	US-10-238-283-91 Sequence 91, Appl
10	264	13.9	2597	9	US-10-238-370-91 Sequence 91, Appl
11	264	13.9	2597	9	US-10-245-055-91 Sequence 91, Appl
12	264	13.9	2597	9	US-10-245-147-91 Sequence 91, Appl
13	264	13.9	2597	9	US-10-245-730-91 Sequence 91, Appl
14	264	13.9	2597	9	US-10-245-739-91 Sequence 91, Appl
15	264	13.9	2597	9	US-10-246-210-91 Sequence 91, Appl
16	264	13.9	2597	9	US-10-239-136-91 Sequence 91, Appl
17	264	13.9	2597	9	US-10-243-024-91 Sequence 91, Appl
18	264	13.9	2597	9	US-10-243-409-91 Sequence 91, Appl
19	264	13.9	2597	9	US-10-245-033-91 Sequence 91, Appl
20	264	13.9	2597	9	US-10-245-621-91 Sequence 91, Appl
21	264	13.9	2597	9	US-10-243-880-91 Sequence 91, Appl
22	264	13.9	2597	9	US-10-243-085-91 Sequence 91, Appl
23	264	13.9	2597	9	US-10-245-185-91 Sequence 91, Appl
24	264	13.9	2597	9	US-10-245-427-91 Sequence 91, Appl
25	264	13.9	2597	9	US-10-245-473-91 Sequence 91, Appl
26	264	13.9	2597	9	US-10-245-770-91 Sequence 91, Appl
27	264	13.9	2597	9	US-10-245-877-91 Sequence 91, Appl
28	264	13.9	2597	9	US-10-246-976-91 Sequence 91, Appl
29	264	13.9	2597	9	US-10-243-320-91 Sequence 91, Appl
30	264	13.9	2597	9	US-10-242-743-91 Sequence 91, Appl
31	264	13.9	2597	9	US-10-242-845-91 Sequence 91, Appl
32	264	13.9	2597	9	US-10-237-636-91 Sequence 91, Appl
33	264	13.9	2597	9	US-10-238-325-91 Sequence 91, Appl
34	264	13.9	2597	9	US-10-238-346-91 Sequence 91, Appl
35	264	13.9	2597	9	US-10-238-411-91 Sequence 91, Appl
36	264	13.9	2597	9	US-10-243-124-91 Sequence 91, Appl
37	264	13.9	2597	9	US-10-243-425-91 Sequence 91, Appl
38	264	13.9	2597	9	US-10-243-446-91 Sequence 91, Appl
39	264	13.9	2597	9	US-10-245-874-91 Sequence 91, Appl
40	264	13.9	2597	9	US-10-242-653-91 Sequence 91, Appl
41	264	13.9	2597	9	US-10-243-167-91 Sequence 91, Appl
42	264	13.9	2597	9	US-10-243-388-91 Sequence 91, Appl
43	264	13.9	2597	9	US-10-244-947-91 Sequence 91, Appl
44	264	13.9	2597	9	US-10-244-968-91 Sequence 91, Appl
45	264	13.9	2597	9	US-10-244-990-91 Sequence 91, Appl

ALIGNMENTS

RESULT 1  
US-10-245-103-91  
; Sequence 91, Application US/10245103  
; Publication No. US20030068778A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3630R1C112  
; CURRENT APPLICATION NUMBER: US/10/245,103  
; CURRENT FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24



US-09-964-956-13 (1-1896) x US-10-245-107-91 (1-2597)

194	QY	GlutyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet	211
2	Db	GAGTATTTTCCCAACCATCTCCAGCGGAAACTGACCAAGAAGTCTGAGCGGATGGCATG	61
214	QY	PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr	233
62	Db	TTCCGGTACGTCTCCATGATGAGTCTGGCGCTCGATGATTAAGATCCCTTCGGACACC	121
234	QY	PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe	253
122	Db	TTCAACATCATCCCTGACTTTGATATCTACTATGTCATGGTTTTAGCAGTGGCAACTTT	181
254	QY	ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu	273
182	Db	GTCTACTTTTTGACCCCTCAACCTGAGATGTGTCTCCACCAGGCTCCACCACCAAGGAG	241
274	QY	GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr	293
242	Db	CAGGTGTATACATCCAGCTCTGAGGCTTTGCAAGGAGGACACAGCCTTCAACTCCCTAT	301
294	QY	ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla	313
302	Db	GTAGAGTGCCCATTTGGCTGTGAGCGCAGTGGGTGGAGTACCGGCTCTGTCAGGCTGCC	361
314	QY	TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp	333
362	Db	TACTGTCCAAAGCGGGGGCGTCTTGGCAGGACCCCTGGAGTCCATCCAGATGATGAC	421
334	QY	LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer	353
422	Db	CTGCTCTTACCGCTTCTTCCAAGGGCCAGAGCGGAANAATGAAATCCCTGGATGATCG	481
354	QY	AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer	373
482	Db	GCCTGTGCATCTTCATCTTGAAGCAGATAAATGACCGCATTAAGGAGCGGTGAGTCT	541
374	QY	CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys	393
542	Db	TGTTACCGGGCGAGGGACGCTGGACCTGGCCTGGCTCAAGGTGAAGGACATCCCTCG	601
394	QY	SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu	413
602	Db	AGCAGTCGCTCTTAACCATTGACGATAACTTCTGTGGCCTGGACATGAATCTCCCTCG	661
414	QY	GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr	433
662	Db	GGAGTGTCCGACATGGTGGCTGGAATTTCCCGTCTTTCACGGAGGACAGGACCGCATG	721
434	QY	SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly	453
722	Db	TCGTGTCATCGATATGTCTACAGAACCACTCTCTGGGCTTTGTGGGACCAAAAGTGGC	781
454	QY	LysLeuLysLys	457
782	Db	AAGCTGAAGAAG	793

### RESULT 3

```

US-10-245-143-91
; Sequence 91, Application US/10245143
; Sequence No. US20030068780A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Wacanze, Colin
; APPLICANT: Wood, William

```

Qy	194	GltTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet	213
Dd	2	GAGTATTTCCCAACCATCTCAGCGGAAACTGACCAAGAACTCTGAGGCGGATGGCATG	61
Qy	214	PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr	233
Dd	62	TTCGCGTAGCTTCCATGATGAGTCTGTGGCCTCGATGATTAAAGATCCCTTCGGACACC	121
Qy	234	PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe	253
Dd	122	TTCAACATCATCCCTGACTTTTGATATCTACTATGCTATGGTTTTAGCAGTGGCAACTTT	181
			273

; SEQ ID NO 91

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```
; Sequence 91, Application US/10245851
; Publication No. US20030068782A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C93
; CURRENT APPLICATION NUMBER: US/10/245,851
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 91
; LENGTH: 2597
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-851-91

Alignment Scores:
Pred. No.: 9,37e-268 Length: 2597
Score: 264.00 Matches: 264
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.92% Indels: 0
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x US-10-245-851-91 (1-2597)

Qy 194 GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
Db 2 GAGTATTTCCCACTCTCCAGCCGGAAACTGACCAAGAACTCTGAGCGCGATGGCATG 61

Qy 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233
Db 62 TTCGCGTACGCTCTCCATGATGAGTTCGTCGCTCGATGATTAGATCCCTTCGGACACC 121

Qy 234 PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253
Db 122 TTCACCATCATCCCTGACTTGTATCTACTATGCTATGCTTGTATGCTTGTACGAGTGGCAACTT 181

Qy 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
Db 182 GTCTACTTTTGGCCCTCCAACTGAGATGGTGTCTCCACGAGGCTCCACCAAGGAG 241

; Sequence 91, Application US/10245851
; Publication No. US20030068782A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C93
; CURRENT APPLICATION NUMBER: US/10/245,851
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 91
; LENGTH: 2597
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-851-91

Qy 274 GlnValTyrThrSerLysLeuValArgLeuLeuCysLysGluAspThrAlaPheAsnSerTyr 293
Db 242 CAGGTGTATACATCCAAAGCTCTGTGAGGCTTTTCAAGGAGGACACAGCCCTTCAACTCCTAT 301

Qy 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313
Db 302 GTAGAGGTGCCCATTTGGCTGTGAGCGCAGTGGGGTGGAGTACCGCTGCTGCAGGCTGCC 361

Qy 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333
Db 362 TACCTGTCTCAAAGCGGGGCGCTGTGCTGGCAGGACCCCTTGGAGTCCATCCAGATGATGAC 421

Qy 334 LeuLeuPheThrValPheSerLysGlnLysArgLysMetLysSerLeuAspGluSer 353
Db 422 CTGCTCTTCAACGCTCTTCTCCAAGGCCAGAACCGGAAATGAATCCCTGGATGAGTGC 481

Qy 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
Db 482 GCCCTGTGCATCTTCAATCTTGAAGCAGATAAATGACCGCATTAAGGAGCGGCTGCAGTCT 541

Qy 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
Db 542 TGTACCAGGGCGAGGCGCAGCTGGCCTGGCTCAAGGTGAAGGACATCCCTCTGC 601

Qy 394 SerSerAlaLeuLeuThrIleAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413
Db 602 AGCAGTGGCTCTTAAACCATTTGACGATAACTTCTGTGGCTGGACATGAATGCTCCCTG 661

Qy 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433
Db 662 GGAGTGTCCGACATGGTGGTGAATTCCTCGCTTTCACGGAGGACAGGACCGCATGAGC 721

Qy 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453
Db 722 TCTGTCTATCGCATATGTCTACAAGAACCACTCTCTGGCCTTTTGTGGGCAACAAAGTGCC 781

Qy 454 LysLeuLysLys 457
Db 782 AAGCTGAAGAG 793

RESULT 6
US-10-245-883-91
; Sequence 91, Application US/10245883
; Publication No. US20030068783A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C70
; CURRENT APPLICATION NUMBER: US/10/245,883
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
```

Mon Jun 30 08:37:45 2003

PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 91
LENGTH: 2597
TYPE: DNA
ORGANISM: Homo Sapien
US-10-245-883-91

Alignment Scores:
Pred. No.: 9,37e-268 Length: 2597
Score: 264.00 Matches: 264
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.92% Indels: 0
DB: Gaps: 0

US-09-964-956-13 (1-1896) x US-10-245-883-91 (1-2597)
QY 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
DB 2 GAGTATTTCCCACTCCACCGGAACTCCACCAAGAACTCTGAGGCGGATGGCATG 61
QY 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233
DB 62 TTCGCGTACGCTTCCATGATGATGTCGTGCGCTCGATGATTAGATCCCTTCGGACACC 121
QY 234 PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253
DB 122 TTCACCATCATCCTGACTTTGATATCTACTATGCTATGCTTTTAGCAGTGGCACTTT 181
QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
DB 182 GTCTACTTTTGGACCTTCCAACTGAGATGGTGTCTCCACGAGCTCCACCAAGGAG 241
QY 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293
DB 242 CAGGTGTATACATCAAGCTCTGAGGCTTTGCAAGGAGGACACAGCCTTCAACTCTAT 301
QY 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuGlnAlaAla 313
DB 302 GTACAGGTGCCCATTTGGCTGTGAGCGCAGTGGGTGGAGTACCGCTGCTGAGGCTGCC 361
QY 314 TyrLeuSerLysAlaGlyAlaValIleuGlyArgThrLeuGlyValHisProAspAsp 333
DB 362 TACCTGTCCAAGCGGGGCGCTGCTTGGCAGGACCTTGGAGTCCATCCAGATGATGAC 421
QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353
DB 422 CTGCTCTTCCAGCTTCTTCCAAAGGCCAGAGCGGAAATGAAATCCCTGGATGAGTGG 481
QY 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
DB 482 GCCCTGTGCATCTTCTATCTTGAAGCAGATAAATGACCGCATTAAGAGCGGCTGCAGTCT 541
QY 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
DB 542 TGTACCGGGGAGGCGACGCTGGACCTGGCTGGCTCAAGTCAAGTGAGGACATCCCTGC 601
QY 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413
DB 602 AGCAGTGGCCTCTTAACCATTCAGGATAACTTCTGTGGCTGGCAGATGATGCTCCCTG 661
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433

Db 662 GGAGTGTCCGACATGCTGCGTGAATTCCTCCGTCCTCACGAGGACAGGGACCGCATGACG 721
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453
Db 722 TCTGTTCATGCATATGCTGTACAGAACCACCTCTCTGCGCTTTGTGGCACCACAAAGTGGC 781
QY 454 LysLeuLysLys 457
Db 782 AAGCTGAAGAAG 793

RESULT 7
US-10-237-535-91
Sequence 91, Application US/10237535
Publication No. US20030073188A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Phillippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3630R1C3
CURRENT APPLICATION NUMBER: US/10/237,535
CURRENT FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27

; PRIOR APPLICATION NUMBER: 60/133459  
 ; PRIOR FILING DATE: 1999-05-11  
 ; PRIOR APPLICATION NUMBER: 60/135725  
 ; PRIOR FILING DATE: 1999-05-25  
 ; PRIOR APPLICATION NUMBER: 60/135729  
 ; PRIOR FILING DATE: 1999-05-25  
 ; PRIOR APPLICATION NUMBER: 60/135750  
 ; PRIOR FILING DATE: 1999-05-25  
 ; PRIOR APPLICATION NUMBER: 60/138385  
 ; PRIOR FILING DATE: 1999-06-09  
 ; PRIOR APPLICATION NUMBER: 60/140653  
 ; PRIOR FILING DATE: 1999-06-22  
 ; PRIOR APPLICATION NUMBER: 60/141037  
 ; PRIOR FILING DATE: 1999-06-23  
 ; PRIOR APPLICATION NUMBER: 60/144732  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/144758  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/144790  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/145228  
 ; PRIOR FILING DATE: 1999-07-20  
 ; PRIOR APPLICATION NUMBER: 60/145698  
 ; PRIOR FILING DATE: 1999-07-26  
 ; PRIOR APPLICATION NUMBER: 60/146222  
 ; PRIOR FILING DATE: 1999-07-28  
 ; PRIOR APPLICATION NUMBER: 60/146843  
 ; PRIOR FILING DATE: 1999-08-03  
 ; PRIOR APPLICATION NUMBER: 60/148188  
 ; PRIOR FILING DATE: 1999-08-10  
 ; PRIOR APPLICATION NUMBER: 60/148513  
 ; PRIOR FILING DATE: 1999-08-12  
 ; PRIOR APPLICATION NUMBER: 60/149327  
 ; PRIOR FILING DATE: 1999-08-17  
 ; PRIOR APPLICATION NUMBER: 60/149395  
 ; PRIOR FILING DATE: 1999-08-17  
 ; PRIOR APPLICATION NUMBER: 60/150114  
 ; PRIOR FILING DATE: 1999-08-20  
 ; PRIOR APPLICATION NUMBER: 60/151700  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: 60/151734  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: 60/152506  
 ; PRIOR FILING DATE: 1999-10-29  
 ; PRIOR APPLICATION NUMBER: 60/170262  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 60/177118  
 ; PRIOR FILING DATE: 2000-01-20  
 ; PRIOR APPLICATION NUMBER: 60/179851  
 ; PRIOR FILING DATE: 2000-02-02  
 ; PRIOR APPLICATION NUMBER: 60/180921  
 ; PRIOR FILING DATE: 2000-02-08  
 ; PRIOR APPLICATION NUMBER: 60/187202  
 ; PRIOR FILING DATE: 2000-03-03  
 ; PRIOR APPLICATION NUMBER: 60/198587  
 ; PRIOR FILING DATE: 2000-04-18  
 ; PRIOR APPLICATION NUMBER: 60/199614  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 60/206330  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/206368  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/209832  
 ; PRIOR FILING DATE: 2000-06-05  
 ; PRIOR APPLICATION NUMBER: 60/218371  
 ; PRIOR FILING DATE: 2000-07-13  
 ; PRIOR APPLICATION NUMBER: 60/222695  
 ; PRIOR FILING DATE: 2000-08-02  
 ; PRIOR APPLICATION NUMBER: 60/229896  
 ; PRIOR FILING DATE: 2000-09-01  
 ; PRIOR APPLICATION NUMBER: 60/230621  
 ; PRIOR FILING DATE: 2000-09-05  
 ; PRIOR APPLICATION NUMBER: 60/232887

; PRIOR FILING DATE: 2000-09-15  
 ; PRIOR APPLICATION NUMBER: 60/235147  
 ; PRIOR FILING DATE: 2000-09-22  
 ; PRIOR APPLICATION NUMBER: 60/261878  
 ; PRIOR FILING DATE: 2001-01-12  
 ; PRIOR APPLICATION NUMBER: 60/261910  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/261939  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/262150  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/264395  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 60/266421  
 ; PRIOR FILING DATE: 2001-02-02  
 ; PRIOR APPLICATION NUMBER: 60/267623  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/274399  
 ; PRIOR FILING DATE: 2001-03-09  
 ; PRIOR APPLICATION NUMBER: 60/280982  
 ; PRIOR FILING DATE: 2001-04-03  
 ; PRIOR APPLICATION NUMBER: 60/282129  
 ; PRIOR FILING DATE: 2001-04-04  
 ; PRIOR APPLICATION NUMBER: 60/282199  
 ; PRIOR FILING DATE: 2001-04-04  
 ; PRIOR APPLICATION NUMBER: 60/290589  
 ; PRIOR FILING DATE: 2001-05-09  
 ; PRIOR APPLICATION NUMBER: 09/180997  
 ; PRIOR FILING DATE: 1998-11-19  
 ; PRIOR APPLICATION NUMBER: 09/267213  
 ; PRIOR FILING DATE: 1999-03-12  
 ; PRIOR APPLICATION NUMBER: 09/380137  
 ; PRIOR FILING DATE: 1999-08-25  
 ; PRIOR APPLICATION NUMBER: 09/380138  
 ; PRIOR FILING DATE: 1999-08-25  
 ; PRIOR APPLICATION NUMBER: 09/403297  
 ; PRIOR FILING DATE: 1999-10-18  
 ; PRIOR APPLICATION NUMBER: 09/423741  
 ; PRIOR FILING DATE: 1999-11-10  
 ; PRIOR APPLICATION NUMBER: 09/709238  
 ; PRIOR FILING DATE: 2000-11-08  
 ; PRIOR APPLICATION NUMBER: 09/802706  
 ; PRIOR FILING DATE: 2001-03-09  
 ; PRIOR APPLICATION NUMBER: 09/872035  
 ; PRIOR FILING DATE: 2001-06-01  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 09/924419  
 ; PRIOR FILING DATE: 2001-08-06  
 ; PRIOR APPLICATION NUMBER: 09/927796  
 ; PRIOR FILING DATE: 2001-08-09  
 ; PRIOR APPLICATION NUMBER: 09/929404  
 ; PRIOR FILING DATE: 2001-08-13  
 ; PRIOR APPLICATION NUMBER: 09/931836  
 ; PRIOR FILING DATE: 2001-08-16  
 ; PRIOR APPLICATION NUMBER: 09/941992  
 ; PRIOR FILING DATE: 2001-08-28  
 ; PRIOR APPLICATION NUMBER: 09/946374  
 ; PRIOR FILING DATE: 2001-09-04  
 ; PRIOR APPLICATION NUMBER: 10/001054  
 ; PRIOR FILING DATE: 2001-11-30  
 ; PRIOR APPLICATION NUMBER: 10/052586  
 ; PRIOR FILING DATE: 2002-01-15  
 ; PRIOR APPLICATION NUMBER: 10/081056  
 ; PRIOR FILING DATE: 2002-02-20  
 ; PRIOR APPLICATION NUMBER: 10/119480  
 ; PRIOR FILING DATE: 2002-04-09

Alignment Scores:

Pred. No.:	9, 37e-268	Length:	2597
Score:	264.00	Matches:	264
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

```

Query Match:      13.92%      Indels:      0
DB:              9          Gaps:      0

US-09-964-956-13 (1-1896) x US-10-237-535-91 (1-2597)

Qy 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
Db 2  GAGTATTTCACCACTCCAGCGGAACTGACCAAGAACTGAGCGGATGGCATG 61
Qy 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233
Db 62 TTCGCGTACGTCCTCCATGATGAGTTCGTGCGCTCGATGATTAAGATCCCTTCGGACACC 121
Qy 234 PheThrIleLeuProAspPheAspIleTyrValTyrGlyPheSerSerGlyAsnPhe 253
Db 122 TTCACCATCATCCCTGACTTTGATATCATCTATGCTATGTTTATGAGTGGCAACTTT 181
Qy 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProProGlySerThrThrLysGlu 273
Db 182 GTCTACTTTTGGACCTCCCAACCTGAGATGGTGTCTCCACCAAGGCTCCACCAAGGAG 241
Qy 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293
Db 242 CAGGTGTATACATCCCAAGCTGCTGAGGCTTTGCAAGAGGAGACACAGCCTTCAACTCCAT 301
Qy 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313
Db 302 GTAGAGTGGCCATTGGCTGTGAGCGCAGTGGGTGGAGTACCGGCTGCTGCAGGCTGCC 361
Qy 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333
Db 362 TACCTGTCCAAAGCGGGGGCGTGTGGCAGGACCCCTTGGAGTCCATCCAGATGATGAC 421
Qy 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353
Db 422 CTGCTCTTCCACCGCTTCTCCAAAGGGCCAGAGCGGAAATGAAATCCTTGATGAGTCG 481
Qy 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
Db 482 GCCCTGTGCATCTTCATCTTGAAGCAGATAAATGACCGCATTAAGAGCGGCTGCAGTCT 541
Qy 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
Db 542 TGTACCGGGGCGAGGGCAGCGTGGACCTGGCTGCTCAAGGTGAAGGACATCCCTGTC 601
Qy 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413
Db 602 AGCAGTGGCTCTTAACCATTTGACCATTAACCTCTGTCGCTGGACATGAATGCTCCCTG 661
Qy 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433
Db 662 GGAGTGTCCGACATGGTGGTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
Qy 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValClyThrLysSerGly 453
Db 722 TCTGTCTCGCATATGTCTACAAAGAACCACTCTCTGGCCCTTTGTGGGCAACCAAGTGGC 781
Qy 454 LysLeuLysLys 457
Db 782 RAGCTGAAGAG 793

RESULT 8
US-10-238-183-91
; Sequence 91, Application US/10238183
; Publication NO. US20030073189A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria

```

```

; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanbe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C11
; CURRENT APPLICATION NUMBER: US/10/238,183
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091358
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/106932
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/115554
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119342
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/123957
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123972
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/127372
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/131271
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/133459
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/135725
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135729
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138385
; PRIOR FILING DATE: 1999-06-09
; PRIOR APPLICATION NUMBER: 60/140653
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144732
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144790
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145228
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698

```

1	PRIOR FILING DATE: 1999-07-26	2
2	PRIOR APPLICATION NUMBER: 60/146222	3
3	PRIOR FILING DATE: 1999-07-28	4
4	PRIOR APPLICATION NUMBER: 60/146843	5
5	PRIOR FILING DATE: 1999-08-03	6
6	PRIOR APPLICATION NUMBER: 60/148188	7
7	PRIOR FILING DATE: 1999-08-10	8
8	PRIOR APPLICATION NUMBER: 60/148513	9
9	PRIOR FILING DATE: 1999-08-12	10
10	PRIOR APPLICATION NUMBER: 60/149327	11
11	PRIOR FILING DATE: 1999-08-17	12
12	PRIOR APPLICATION NUMBER: 60/149395	13
13	PRIOR FILING DATE: 1999-08-17	14
14	PRIOR APPLICATION NUMBER: 60/150114	15
15	PRIOR FILING DATE: 1999-08-20	16
16	PRIOR APPLICATION NUMBER: 60/151700	17
17	PRIOR FILING DATE: 1999-08-31	18
18	PRIOR APPLICATION NUMBER: 60/151734	19
19	PRIOR FILING DATE: 1999-08-31	20
20	PRIOR APPLICATION NUMBER: 60/162506	21
21	PRIOR FILING DATE: 1999-10-29	22
22	PRIOR APPLICATION NUMBER: 60/170262	23
23	PRIOR FILING DATE: 1999-12-09	24
24	PRIOR APPLICATION NUMBER: 60/177118	25
25	PRIOR FILING DATE: 2000-01-20	26
26	PRIOR APPLICATION NUMBER: 60/179851	27
27	PRIOR FILING DATE: 2000-02-02	28
28	PRIOR APPLICATION NUMBER: 60/180921	29
29	PRIOR FILING DATE: 2000-02-08	30
30	PRIOR APPLICATION NUMBER: 60/187202	31
31	PRIOR FILING DATE: 2000-03-03	32
32	PRIOR APPLICATION NUMBER: 60/198587	33
33	PRIOR FILING DATE: 2000-04-18	34
34	PRIOR APPLICATION NUMBER: 60/199614	35
35	PRIOR FILING DATE: 2000-04-25	36
36	PRIOR APPLICATION NUMBER: 60/206330	37
37	PRIOR FILING DATE: 2000-05-23	38
38	PRIOR APPLICATION NUMBER: 60/206368	39
39	PRIOR FILING DATE: 2000-05-23	40
40	PRIOR APPLICATION NUMBER: 60/209832	41
41	PRIOR FILING DATE: 2000-06-05	42
42	PRIOR APPLICATION NUMBER: 60/218371	43
43	PRIOR FILING DATE: 2000-07-13	44
44	PRIOR APPLICATION NUMBER: 60/222695	45
45	PRIOR FILING DATE: 2000-08-02	46
46	PRIOR APPLICATION NUMBER: 60/229896	47
47	PRIOR FILING DATE: 2000-09-01	48
48	PRIOR APPLICATION NUMBER: 60/230621	49
49	PRIOR FILING DATE: 2000-09-05	50
50	PRIOR APPLICATION NUMBER: 60/232887	51
51	PRIOR FILING DATE: 2000-09-15	52
52	PRIOR APPLICATION NUMBER: 60/235147	53
53	PRIOR FILING DATE: 2000-09-22	54
54	PRIOR APPLICATION NUMBER: 60/261878	55
55	PRIOR FILING DATE: 2001-01-12	56
56	PRIOR APPLICATION NUMBER: 60/261910	57
57	PRIOR FILING DATE: 2001-01-16	58
58	PRIOR APPLICATION NUMBER: 60/261939	59
59	PRIOR FILING DATE: 2001-01-16	60
60	PRIOR APPLICATION NUMBER: 60/262150	61
61	PRIOR FILING DATE: 2001-01-16	62
62	PRIOR APPLICATION NUMBER: 60/264395	63
63	PRIOR FILING DATE: 2001-01-25	64
64	PRIOR APPLICATION NUMBER: 60/266421	65
65	PRIOR FILING DATE: 2001-02-02	66
66	PRIOR APPLICATION NUMBER: 60/267623	67
67	PRIOR FILING DATE: 2001-02-09	68
68	PRIOR APPLICATION NUMBER: 60/274399	69
69	PRIOR FILING DATE: 2001-03-09	70
70	PRIOR APPLICATION NUMBER: 60/280982	71
71	PRIOR FILING DATE: 2001-04-03	72
72	PRIOR APPLICATION NUMBER: 60/282129	73
73	PRIOR FILING DATE: 2001-04-04	74

1	PRIOR APPLICATION NUMBER: 60/282199
2	PRIOR FILING DATE: 2001-04-04
3	PRIOR APPLICATION NUMBER: 60/290589
4	PRIOR FILING DATE: 2001-05-09
5	PRIOR APPLICATION NUMBER: 09/180997
6	PRIOR FILING DATE: 1998-11-19
7	PRIOR APPLICATION NUMBER: 09/267213
8	PRIOR FILING DATE: 1999-03-12
9	PRIOR APPLICATION NUMBER: 09/380137
10	PRIOR FILING DATE: 1999-08-25
11	PRIOR APPLICATION NUMBER: 09/380138
12	PRIOR FILING DATE: 1999-08-25
13	PRIOR APPLICATION NUMBER: 09/403297
14	PRIOR FILING DATE: 1999-10-18
15	PRIOR APPLICATION NUMBER: 09/423741
16	PRIOR FILING DATE: 1999-11-10
17	PRIOR APPLICATION NUMBER: 09/7709238
18	PRIOR FILING DATE: 2000-11-08
19	PRIOR APPLICATION NUMBER: 09/802706
20	PRIOR FILING DATE: 2001-03-09
21	PRIOR APPLICATION NUMBER: 09/872035
22	PRIOR FILING DATE: 2001-06-01
23	PRIOR APPLICATION NUMBER: 09/918585
24	PRIOR FILING DATE: 2001-07-30
25	PRIOR APPLICATION NUMBER: 09/924419
26	PRIOR FILING DATE: 2001-08-06
27	PRIOR APPLICATION NUMBER: 09/927796
28	PRIOR FILING DATE: 2001-08-09
29	PRIOR APPLICATION NUMBER: 09/929404
30	PRIOR FILING DATE: 2001-08-13
31	PRIOR APPLICATION NUMBER: 09/931836
32	PRIOR FILING DATE: 2001-08-16
33	PRIOR APPLICATION NUMBER: 09/941992
34	PRIOR FILING DATE: 2001-08-28
35	PRIOR APPLICATION NUMBER: 09/946374
36	PRIOR FILING DATE: 2001-09-04
37	PRIOR APPLICATION NUMBER: 10/001054
38	PRIOR FILING DATE: 2001-11-30
39	PRIOR APPLICATION NUMBER: 10/052586
40	PRIOR FILING DATE: 2002-01-15
41	PRIOR APPLICATION NUMBER: 10/081056
42	PRIOR FILING DATE: 2002-02-20
43	PRIOR APPLICATION NUMBER: 10/119480
44	PRIOR FILING DATE: 2002-04-09

Alignment Scores:

Alignment Scores:		
Pred. No.:	9.37e-268	Length: 2597
Score:	264.00	Matches: 264
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	13.92%	Indels: 0
DB:	9	Gaps: 0

US-09-964-956-13 (1-1896) x US-10-238-183-91 (1-2597)

Qy	194	GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet	213
Db	2	GAGTATTTTCCACCATCTCCAGCGGAAACTGACCAAGACTCTGAGCGGATGGCATG	61
Qy	214	PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysileProSerAspThr	233
Db	62	TTCCGCTAGCTTCTCCATGATGAGTTCGTGGCCTCGATGATTAAAGATCCCTTCGGACACC	121
Qy	234	PheThrIleIleProAspPheAspIleTyrValTyrGlyPheSerSerGlyAsnPhe	253
Db	122	TTACACCATCATCCCTGACTTTGTATATCTACTATGCTCTATGGTTTATAGCAGTGGCAACTTT	181
Qy	254	ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu	273
Db	182	GTCTACTATTTTGTAGCCCTCCAACTGAGATGGTGTCTCCACCAGGTCTCCACCACCAAGGAG	241
Qy	274	GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr	293



QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
Db 722 TCTGTCATCCGATATGTCTACAAGAACCACTCTCTGGCCTTTGTGGGCACCAAAAGTGGC 781  
QY 454 LysLeuLysLys 457  
Db 782 AAGCTGAAGAAG 793

RESULT 10  
US-10-238-370-91  
; Sequence 91, Application US/10238370  
; Publication No. US20030073191A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3630R1C10  
; CURRENT APPLICATION NUMBER: US/10/238,370  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 10/137942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; Remaining prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 91  
; LENGTH: 2597  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-238-370-91

Alignment Scores:  
Pred. No.: 9,37e-268 Length: 2597  
Score: 264.00 Matches: 264  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.92% Indels: 0  
DB: Gaps: 9

US-09-964-956-13 (1-1896) x US-10-238-370-91 (1-2597)  
QY 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213  
Db 2 GAGTATTTCCCACTCTCCAGCGAAACTGACCAAGAACTCTGAGCGGATGGCATG 61  
QY 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233

Db 62 TTCGGTACGCTTCCATGATGAGTTCGTGGGCTCGATGATTAAAGATCCCTTCGGACACC 121  
QY 234 PheThrIleLeuProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 233  
Db 122 TTCACCATCATCCCTGACTTTTATGATATCTACTATGTCTATGGTWTATAGCAGTGCAACTTT 181  
QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273  
Db 182 GTCCTACTTTTTCACCTCCAACTGAGATGGTGTCTCCACAGGCTCCACCAAGGAG 241  
QY 274 GlnValTyrThrSerLysLeuValArgLeuCysLysLysGluAspThrAlaPheAsnSerTyr 293  
Db 242 CAGGTGTATACATCCAAAGCTCGTGGCTTTCAGAGGAGGACACAGCCCTTCAACTCCTAT 301  
QY 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313  
Db 302 GTAGAGGTGCCCTTGGCTGTGAGCGCAGTGGGTGGAGTACCGCTGTCTGCAAGGCTGCC 361  
QY 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333  
Db 362 TACCTGTCCAAAGCGGGCGCTGTGGCAGGACCCCTTGGAGTCCATCCAGATGATGAC 421  
QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353  
Db 422 CTGCTCTTTCACCGCTCTTCTCAAGGCGCCAGAAAGCGGAAATGAAATCCCTGGATGAGTCG 481  
QY 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373  
Db 482 GCCCTGTGCATCTTCATCTTGAAGCAGATAAATGACCGCATTAAGGAGCGGCTGCCAGTCT 541  
QY 374 CystYrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393  
Db 542 TGTACCAGGCGGAGGCGACGCTGGACCTGGCTGCTCAAGGTGAAGGACATCCCTCGC 601  
QY 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413  
Db 602 AGCAGTGCCTCTTAACCATTTGACGATAAATCTCTGTGGCTGGACATGAATGCTCCCTG 661  
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433  
Db 662 GGAGTGTCCGACATGGTGGTGAATTCCTGCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAG 721  
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
Db 722 TCTGTCAATCGCATATGTCTACAAGAACCACTCTCTGGCCTTTGTGGGCACCAAAAGTGGC 781  
QY 454 LysLeuLysLys 457  
Db 782 AAGCTGAAGAAG 793

RESULT 11  
US-10-245-055-91  
; Sequence 91, Application US/10245055  
; Publication No. US20030073192A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3630R1C88  
; CURRENT APPLICATION NUMBER: US/10/245,055  
; CURRENT FILING DATE: 2002-09-16

;; PRIOR APPLICATION NUMBER: 10/197942  
;; PRIOR FILING DATE: 2002-07-18  
;; PRIOR APPLICATION NUMBER: 60/059114  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/063046  
;; PRIOR FILING DATE: 1997-10-24  
;; PRIOR APPLICATION NUMBER: 60/065027  
;; PRIOR FILING DATE: 1997-11-10  
;; PRIOR APPLICATION NUMBER: 60/079689  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/086478  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087607  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090689  
;; PRIOR FILING DATE: 1998-06-25  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; SEQ ID NO 91  
;; LENGTH: 2597  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-10-245-055-91

Alignment Scores:  
Pred. No.: 9.37e-268 Length: 2597  
Score: 264.00 Matches: 264  
Percent Similarity: 100.00% Conservations: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.92% Indels: 0  
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x US-10-245-055-91 (1-2597)

Qy 194 GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213  
Db 2 GAGTATTTCCACCATCTCCAGCGGAACTGACCAAGAACTGTAGCGGATGCGATG 61  
Qy 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233  
Db 62 TTCGGTACGTCTCCATGATGAGTCTGGCTCGATGATTAAGATCCCTTCGACAC 121  
Qy 234 PheThrIleLeuProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253  
Db 122 TTCACCATCATCCCTGACTTTGATATCTACTATGTCTATGGTTTAGCAGTGGCACTT 181  
Qy 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273  
Db 182 GTCTACTTTTGACCCCTCCACCTCAGATGCTGTCTCCACCAAGGCTCCACCAAGGAG 241  
Qy 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293  
Db 242 CAGGTGTATACATCAAGCTCGTAGGCTTTGCAAGGAGGACACAGCCCTCAACCTCAT 301  
Qy 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaLa 313  
Db 302 GTAGAGGTGCCATTTGGCTGTGAGCGCAGTGGGTGGAGTACCGCTGTCTGAGGCTGCC 361  
Qy 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333  
Db 362 TACCTGTCCAAAGCGGGCGCGTGTGGCAGGACCTTGGAGTCCATCCAGATGATGAC 421  
Qy 334 LeuLeuPheThrValPheSerGlyGlnLysArgLysMetLysSerLeuAspGluSer 353  
Db 422 CTGCTCTTACCGCTCTTCTCCAAAGGCGGAGAGCGGAAATGAAATCCCTGGATGAGTCG 481  
Qy 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373  
Db 482 GCCCTGTGCACTTCTATCTTGAAGCAGATAAATGACCGCATTAAGGAGCGGCTGCAGTCT 541

Qy 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393  
Db 542 TGTACCAGGGGCGAGGCGACGCTGGACCTGGCTCAAGGTGAAGGACATCCCTCGC 601  
Qy 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413  
Db 602 AGCAGTGGCGCTCTTAACCATTTGACGATAACTTCTGTGGCTGGACATGAATGCTCCCTG 661  
Qy 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433  
Db 662 GGAGTGTCCGACATGCTGCTGGAATTCCTGCTTCCAGGAGACACGGGACCATGACG 721  
Qy 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
Db 722 TCTGTGTCATGTCATATGCTTACAAGAACCACTCTCTGCTGCTTGTGGGCACCAAAAGTGGC 781  
Qy 454 LysLeuLysLys 457  
Db 782 AAGCTCAAGAAG 793  
RESULT 12  
US-10-245-147-91  
;; Sequence 91, Application US/10245147  
;; Publication No. US20030073193A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Baker, Kevin  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Stephan, Jean-Phillippe  
;; APPLICANT: Watanabe, Colin  
;; APPLICANT: Wood, William  
;; APPLICANT: Zhang, Zemin  
;; APPLICANT: Fong, Sherman  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3630R1C72  
;; CURRENT APPLICATION NUMBER: US/10/245,147  
;; CURRENT FILING DATE: 2002-09-16  
;; PRIOR APPLICATION NUMBER: 10/197942  
;; PRIOR FILING DATE: 2002-07-18  
;; PRIOR APPLICATION NUMBER: 60/059114  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/063046  
;; PRIOR FILING DATE: 1997-10-24  
;; PRIOR APPLICATION NUMBER: 60/065027  
;; PRIOR FILING DATE: 1997-11-10  
;; PRIOR APPLICATION NUMBER: 60/079689  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/086478  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087607  
;; PRIOR FILING DATE: 1998-06-02  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090689  
;; PRIOR FILING DATE: 1998-06-25  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 116  
;; SEQ ID NO 91  
;; LENGTH: 2597  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-10-245-147-91

Alignment Scores:  
Pred. No.: 9.37e-268 Length: 2597

Score: 264.00 Matches: 264  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.92% Indels: 0  
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x US-10-245-147-91 (1-2597)

```
QY 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
DB 2 GAGTATTTTCCCAACCATCTCCAGCGGAAACTGACCAAGAACTCTGAGCGGATGGCATG 61
QY 214 PheAlaTyrValPheHisaspGluPheValAlaSerMetIleLysIleProSerAspThr 233
DB 62 TTCGCGTACGCTCTCCATGATGAGTTGCGGCTCGATGATTAAGATCCCTTCGGACACC 121
QY 234 PheThrIleIleProaspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253
DB 122 TTCACCATCATCCTGACTTTGATATCTACTATGCTCTATGTTTACAGTGGCAACTTT 181
QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
DB 182 GTCTACTTTTGGACCTCCAACTGAGATGGTGTCTCCACGAGCTCCACCAAGGAG 241
QY 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluaspThrAlaPheAsnSerTyr 293
DB 242 CAGGTGTATACATCCCAAGCTCGTAGGCTTTGCAAGGAGGACACAGCCTTCAACTCTAT 301
QY 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313
DB 302 GTACAGGTGCCATTTGGCTGTGAGCGGATGGGTGGAGTACCCCTGCTCGAGGCTGCC 361
QY 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333
DB 362 TACCTGTCCAAAGCGGGCGCTGTTGGCAGGACCTTGGAGTCCATCCAGATGATGAC 421
QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353
DB 422 CTGCTCTTCCACGCTCTCTCCAAAGGCCAGAGCGGAAATGAAATCCCTGGATGATCG 481
QY 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
DB 482 GCCTGTGCATCTTCATCTTGAAGCAGATAAATGACCGCATTAAGGAGCGGCTCGAGTCT 541
QY 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
DB 542 TGTACCGGGCGGAGGCGACGCTGGACCTGGCTGGCTCAAGGTGAAGGACATCCCTGTC 601
QY 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuaspMetAsnAlaProLeu 413
DB 602 AGCAGTGGCGCTCTTAACCATTTGACGATAAATCTGTGGCTGGACATGAATGCTCCCGTG 661
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluaspArgAspArgMetThr 433
DB 662 GGAGTGTCCGACATGTGCGTGGAAATCCCGTCTTCACGAGGACGAGGACCGCATGACG 721
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453
DB 722 TCTGTATCATCATATGCTACAAAGAACCACTCTCTGGCCCTTTGTGGGCACCAAAAGTGGC 781
QY 454 LysLeuLysLys 457
DB 782 AAGCTGAAGAAG 793
```

RESULT 13

US-10-245-730-91

; Sequence 91, Application US/10245730

; Publication No. US20030073194A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin

; APPLICANT: Eaton, Dan

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Goddard, Audrey

; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P363081C85  
; CURRENT APPLICATION NUMBER: US/10/245,730  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 91  
; LENGTH: 2597  
; TYPE: DNA  
; ORGANISM: Homo Saplen  
US-10-245-730-91

Alignment Scores:

Pred. No.: 9,378-268 Length: 2597  
Score: 264.00 Matches: 264  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.92% Indels: 0  
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x US-10-245-730-91 (1-2597)

```
QY 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
DB 2 GAGTATTTTCCCAACCATCTCCAGCGGAAACTGACCAAGAACTCTGAGCGGATGGCATG 61
QY 214 PheAlaTyrValPheHisaspGluPheValAlaSerMetIleLysIleProSerAspThr 233
DB 62 TTCGCGTACGCTCTCCATGATGAGTTGCGGCTCGATGATTAAGATCCCTTCGGACACC 121
QY 234 PheThrIleIleProaspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253
DB 122 TTCACCATCATCCTGACTTTGATATCTACTATGCTCTATGTTTACAGTGGCAACTTT 181
QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
DB 182 GTCTACTTTTGGACCTCCAACTGAGATGGTGTCTCCACGAGCTCCACCAAGGAG 241
QY 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluaspThrAlaPheAsnSerTyr 293
DB 242 CAGGTGTATACATCCCAAGCTCGTAGGCTTTGCAAGGAGGACACAGCCTTCAACTCTAT 301
QY 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313
DB 302 GTACAGGTGCCATTTGGCTGTGAGCGGATGGGTGGAGTACCCCTGCTCGAGGCTGCC 361
QY 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333
DB 362 TACCTGTCCAAAGCGGGCGCTGTTGGCAGGACCTTGGAGTCCATCCAGATGATGAC 421
QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353
DB 422 CTGCTCTTCCACGCTCTCTCCAAAGGCCAGAGCGGAAATGAAATCCCTGGATGATCG 481
QY 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
DB 482 GCCTGTGCATCTTCATCTTGAAGCAGATAAATGACCGCATTAAGGAGCGGCTCGAGTCT 541
QY 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
DB 542 TGTACCGGGCGGAGGCGACGCTGGACCTGGCTGGCTCAAGGTGAAGGACATCCCTGTC 601
QY 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuaspMetAsnAlaProLeu 413
DB 602 AGCAGTGGCGCTCTTAACCATTTGACGATAAATCTGTGGCTGGACATGAATGCTCCCGTG 661
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluaspArgAspArgMetThr 433
DB 662 GGAGTGTCCGACATGTGCGTGGAAATCCCGTCTTCACGAGGACGAGGACCGCATGACG 721
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453
DB 722 TCTGTATCATCATATGCTACAAAGAACCACTCTCTGGCCCTTTGTGGGCACCAAAAGTGGC 781
QY 454 LysLeuLysLys 457
DB 782 AAGCTGAAGAAG 793
```

Db 302 GTAGAGTGCCCATTTGGCTGTGAGCGCAGTGGGTGGAGTACCGCCTGCTGCAGGCTGCC 361  
Qy 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333  
Db 362 TACCTGTCCAAAGCGGGCGCTGTTGGCAGGACCTTGGAGTCCATCCAGATGATGAC 421  
Qy 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353  
Db 422 CTGCTCTTCCAGCTTCTTCCAAAGCGCAGAGCGGAAATGAATCCCTGGATGAGTCG 481  
Qy 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373  
Db 482 GCCCTGTGCATCTTCATCTTCAAGCAGATAAATGACCGCATTAAGGAGCGGTGCGTCT 541  
Qy 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393  
Db 542 TGTACCGGGCGAGGCGACCTGACCTGGCTGCGCTCAAGGTGAAGGACATCCCTGTC 601  
Qy 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413  
Db 602 AGCAGTGGCTCTTAACCATTTGACGATAAATCTGTGGCTGGACATGAATGCTCCCTG 661  
Qy 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433  
Db 662 GGAGTGTCCGACATGCTGCTGGAATTCCTGCTTACGGGAGGACAGGACCGCATGAGC 721  
Qy 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
Db 722 TCTGTGATCGCATGTGTACAAGAACCACTCTCTGGCTTGTGGGACCAAAAAGTGCC 781  
Qy 454 LysLeuLysLys 457  
Db 782 AAGTGAAGAAG 793

RESULT 14

US-10-245-739-91  
: Sequence 91, Application US/10245739  
: Publication No. US20030073195A1  
: GENERAL INFORMATION:  
: APPLICANT: Baker, Kevin  
: APPLICANT: Eaton, Dan  
: APPLICANT: Filvaroff, Ellen  
: APPLICANT: Goddard, Audrey  
: APPLICANT: Grimaldi, J. Christopher  
: APPLICANT: Gurney, Austin  
: APPLICANT: Smith, Victoria  
: APPLICANT: Stephan, Jean-Philippe  
: APPLICANT: Watanabe, Colin  
: APPLICANT: Wood, William  
: APPLICANT: Zhang, Zemin  
: APPLICANT: Fong, Sherman  
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
: FILE REFERENCE: P3630RLC86  
: CURRENT APPLICATION NUMBER: US/10/245,739  
: CURRENT FILING DATE: 2002-09-16  
: PRIOR APPLICATION NUMBER: 10/197942  
: PRIOR FILING DATE: 2002-07-18  
: PRIOR APPLICATION NUMBER: 60/059114  
: PRIOR FILING DATE: 1997-09-17  
: PRIOR APPLICATION NUMBER: 60/063046  
: PRIOR FILING DATE: 1997-10-24  
: PRIOR APPLICATION NUMBER: 60/065027  
: PRIOR FILING DATE: 1997-11-10  
: PRIOR APPLICATION NUMBER: 60/079689  
: PRIOR FILING DATE: 1998-03-27  
: PRIOR APPLICATION NUMBER: 60/086478  
: PRIOR FILING DATE: 1998-05-22  
: PRIOR APPLICATION NUMBER: 60/087607  
: PRIOR FILING DATE: 1998-06-02  
: PRIOR APPLICATION NUMBER: 60/089801  
: PRIOR FILING DATE: 1998-06-18  
: PRIOR APPLICATION NUMBER: 60/090557

: PRIOR FILING DATE: 1998-06-24  
: PRIOR APPLICATION NUMBER: 60/090689  
: PRIOR FILING DATE: 1998-06-25  
: Remaining Prior Application data removed - See File Wrapper or PALM.  
: NUMBER OF SEQ ID NOS: 116  
: SEQ ID NO 91  
: LENGTH: 2597  
: TYPE: DNA  
: ORGANISM: Homo Sapien  
US-10-245-739-91

Alignment Scores: 9.37e-268 , Length: 2597  
Pred. No.: 264.00 Matches: 264  
Score: 264.00  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.92% Indels: 0  
DB: Gaps: 0  
US-09-964-956-13 (1-1896) x US-10-245-739-91 (1-2597)  
Qy 194 GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213  
Db 2 GAGTATTTTCCCACTCTCCAGCGGAACTGACCAAGAACTCTGAGCGGATGGCATG 61  
Qy 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233  
Db 62 TTCCGCTACGTCTTCCATGATGATTCGTGCTGCTCGATGATTAAGATCCCTTCGACACC 121  
Qy 234 PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253  
Db 122 TTCACCATCATCTGACCTTTGATATCTACTATGCTATGCTATGCTTTAGCAGTGGCACTT 181  
Qy 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273  
Db 182 GTCTACTTTTGGACCTTCCAACTGAGATGCTGCTCCACGAGCTCCACCAAGAGAG 241  
Qy 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293  
Db 242 CAGGTGTATACATCCAAAGCTCGTGAGCTTTGCAAGGAGGACACAGGCTTCACTCTCTAT 301  
Qy 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuGlnAlaAla 313  
Db 302 GTAGAGTGCCCATTTGGCTGTGAGCGCAGTGGGTGGAGTACCGCTGCTGCAGGCTGCC 361  
Qy 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333  
Db 362 TACCTGTCCAAAGCGGGGCGCTGCTTGGCAGGACCTTGGAGTCCATCCAGATGATGAC 421  
Qy 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353  
Db 422 CTGCTCTTCCAGCTTCTTCCAAAGCGCAGAGCGGAAATGAATCCCTGGATGAGTCG 481  
Qy 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373  
Db 482 GCCCTGTGCATCTTCATCTTGAACAGATATAATGACCGCATTAAGGAGCGGCTCGACTCT 541  
Qy 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393  
Db 542 TGTACCGGGCGAGGCGACCTGACCTGGCTGCGCTCAAGGTGAAGGACATCCCTGTC 601  
Qy 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413  
Db 602 AGCAGTGGCTCTTAACCATTTGACGATAAATCTGTGGCTGGACATGAATGCTCCCTG 661  
Qy 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433  
Db 662 GGAGTGTCCGACATGCTGCTGGAATTCCTGCTTACGGGAGGACAGGACCGCATGAGC 721  
Qy 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
Db 722 TCTGTGATCGCATGTGTACAAGAACCACTCTCTGGCTTGTGGGACCAAAAAGTGCC 781

QY 454 LysLeuLysLys 457  
| | | | | | | | | |  
Db 782 AAGCTGAAGAAG 793

RESULT 15

US-10-246-210-91

; Sequence 91, Application US/10246210  
; Publication No. US20030073196A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin

; APPLICANT: Eaton, Dan

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Goddard, Audrey

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin

; APPLICANT: Smith, Victoria

; APPLICANT: Stephan, Jean-Phillippe

; APPLICANT: Watanabe, Colin

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; APPLICANT: Fong, Sherman

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3630R1C121

; CURRENT APPLICATION NUMBER: US/10/246,210

; PRIOR FILING DATE: 2002-09-18

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/063046

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/065027

; PRIOR FILING DATE: 1997-11-10

; PRIOR APPLICATION NUMBER: 60/079689

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/086478

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/090557

; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090689

; PRIOR FILING DATE: 1998-06-25

; Remaining prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 116

; SEQ ID NO 91

; LENGTH: 2597

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-246-210-91

Alignment Scores:  
Pred. No.: 9,37e-268 Length: 2597  
Score: 264.00 Matches: 264  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.92% Indels: 0  
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x US-10-246-210-91 (1-2597)

QY 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213

Db 2 GAGTATTTCCCACTCTCCAGCGGAACCTGACCAAGAACTCTGAGCGGATGGCATG 61

QY 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233

Db 62 TTCGCGTACGCTCTCCATGATGAGTTCGTGGCCTCGATGATTAAGATCCCTTCGGACACC 121

QY 234 PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253

Db 122 TTCACCATCATCCCTGACTTTTGATATCTACTATGTCTATGGTATTTAGCAGTGCAACTTT 181  
QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273  
| | | | | | | | | |  
Db 182 GTCTACTTTTGGACCTTCAACCTGAGATGGTGTCTCCACCAAGGCTCCACCAAGGAG 241  
QY 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293  
| | | | | | | | | |  
Db 242 CAGGTGTATACATCCCAAGCTCGTAGGCTTTGCAAGGAGGACACAGCCTTCAACTCCTAT 301  
QY 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313  
| | | | | | | | | |  
Db 302 GTAGAGGTGCCCATTTGGCTGTGAGCGCAGTGGGGTGGAGTACCGCTGTCTGCAAGGTCGC 361  
QY 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspAsp 333  
| | | | | | | | | |  
Db 362 TACCTGTCCAAAGCGGGGCGCTGTGGCAGGACCTTGGAGTCCATCCAGATGATGAC 421  
QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353  
| | | | | | | | | |  
Db 422 CTGCTCTTTCACGCTCTTCTCCAAAGGCGCAGAAATGAAATCCCTGGATGAGTCG 481  
QY 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373  
| | | | | | | | | |  
Db 482 GCCCTGTGCATCTTCACTTTGAAGCAGATAAATGACCGCATTAAGAGAGCGGCTGCAGTCT 541  
QY 374 CysTyrArgGlyGlyThrLeuAspLeuAlaTyrLeuLysValLysAspIleProCys 393  
| | | | | | | | | |  
Db 542 TGTATTACCGGGGCGAGGCGACGCTGGACCTGGCTGCTCAAGGTGAAGGACATCCCTCGC 601  
QY 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413  
| | | | | | | | | |  
Db 602 AGCAGTGCCTCTTAACCATTCACGATAACTTCTGTGGCTGGACATGAATGCTCCCTCG 661  
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433  
| | | | | | | | | |  
Db 662 GGAGTGTCCGACATGGTGGTGAATTCCTCCAGGAGGACAGGACCGCATGACG 721  
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
| | | | | | | | | |  
Db 722 TCTGTCTATCGCATATGTCTACAAGAACCACTCTCTGGCCTTTGTGGCACCACCAAGTGGC 781  
QY 454 LysLeuLysLys 457  
| | | | | | | | | |  
Db 782 AAGCTGAAGAAG 793

Search completed: June 29, 2003, 03:19:12  
Job time : 640 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 18:34:02 ; Search time 5335 seconds  
(without alignments)  
5755.697 Million cell updates/sec

Title: US-09-964-956-13  
Perfect score: 1896  
Sequence: 1 MKAMPWNWTCLLSHLLMGM.....OKLAYKLEQVITLMSLDSNK 1896

Scoring table:  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308016

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO\_spool/US09964956/runat\_24062003.124829.7368/app\_query.fasta\_1.2055  
-DB=EST -QFMT=fastap -SUFFIX=oli.rst -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.coi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pt  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09964956@cgn.1.1.4160@runat\_24062003.124829.7368 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_othr.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	199	10.5	627	12	BG086250	BG086250 H3123F01-AL707761 DKF2p686H
2	152	8.0	483	9	AL707761	BM105323 508753 MA
3	138	7.3	567	13	BM105323	BF952116 QV1-NN022
4	120	6.3	512	12	BF952116	BF953019 QV3-NN019
5	109	5.7	564	12	BF953019	BB656076 BB656076
6	105	5.5	612	10	BB656076	BG694413 NISC_i.v01
7	77	4.1	412	12	BG694413	AA077700 7B44B08 C
8	76	4.0	232	9	AA077700	AA077556 7B19F10 C
9	75	4.0	548	10	AA018699	BF953008 QV3-NN019
10	71	3.7	233	9	AA077556	AL595528 AL595528
11	68	3.6	251	9	AA077556	BQ434703 AGENCOURT
12	66	3.5	493	12	BF953008	BQ878090 AGENCOURT
13	63	3.3	650	9	AL595528	BQ684665 AGENCOURT
14	63	3.3	844	14	BQ434703	BQ683335 AGENCOURT
15	63	3.3	851	14	BQ878090	BF312056 601897930
16	63	3.3	920	14	BQ684665	BQ180469 UI-M-EXO-
17	63	3.3	933	14	BQ683335	BQ180472 UI-M-EXO-
18	58	3.1	924	12	BF312056	BQ680140 AGENCOURT
19	57	3.0	662	14	BQ180469	BF725116 bx12d06.Y
20	57	3.0	664	14	BQ180472	BF725116 bx12d06.Y
21	55	2.9	881	14	BQ680140	BG912949 602807353
22	54	2.8	558	14	BM946917	BG912949 602807353
23	54	2.8	585	12	BF725116	BG910147 602805311
24	54	2.8	643	13	BG912949	BG910147 602805311
25	54	2.8	689	13	BQ39160	BG910147 602805311
26	54	2.8	730	12	BG910147	BG910147 602805311
27	54	2.8	777	12	BE910224	BQ897297 AGENCOURT
28	54	2.8	927	14	BQ897297	BQ897297 AGENCOURT
29	54	2.8	960	13	BM561910	BQ897297 AGENCOURT
30	54	2.8	971	13	BI905192	BI905192 603167521
31	54	2.8	976	12	BF527763	BF527763 602040830
32	53	2.8	456	13	BM484361	BM484361 538194 MA
33	52	2.7	360	9	AU245221	AU245221 AU245221
34	51	2.7	242	9	AA077496	AA077496 7B18G11 C
35	51	2.7	351	12	BF388606	BF388606 UI-R-BS2-
36	51	2.7	442	12	BG834999	BG834999 353406 MA
37	50	2.6	789	12	BG216679	BG216679 RST36373
38	49	2.6	513	13	BI512619	BI512619 BB160009B
39	49	2.6	748	12	BF311719	BF311719 601897214
40	49	2.6	972	12	BF304376	BF304376 601887286
41	46	2.4	441	14	HI10550	HI10550 ym04d05.r1
42	46	2.4	464	14	HI15503	HI15503 ym29f06.r1
43	45	2.4	627	10	BB452530	BB452530 BB452530
44	45	2.4	729	13	BJ072104	BJ072104 BJ072104
45	45	2.4	762	9	ALI34739	ALI34739 DKF2p547C

## ALIGNMENTS

RESULT 1  
BG086250  
LOCUS  
DEFINITION  
H3123F01-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
627 bp mRNA linear EST 26-JAN-2001  
H3123F01 5', mRNA sequence.  
ACCESSION  
BG086250  
VERSION  
BG086250.1 GI:12568814  
KEYWORDS  
EST.  
SOURCE  
house mouse  
ORGANISM  
Mus musculus  
REFERENCE  
1 (bases 1 to 627)  
Eukaryota: Metazoa; Chordata; Vertebrata: Euteleostomi;  
Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
AUTHORS  
Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka, T.S., Carter, M.G. and KO, M.S.H.

# TITLE JOURNAL COMMENT

Verification and initial annotation of NTA mouse 15K cDNA clone set  
Unpublished (2001)  
Other ESTs: H3123F01-3  
Contact: George J. Kargul  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
This clone set has been freely distributed to the community. Please  
visit <http://igsun.grc.nia.nih.gov/cdna/15k.html> for details.  
Plate: H3123 Row: F Column: 01  
Seq primer: -21M13 Reverse  
High quality sequence stop: 627  
POLYA-No.

## FEATURES

### Source

1. 627  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="niaEST:H3123F01-5"  
/db\_xref="taxon:10090"  
/clone="H3123F01"  
/clone\_lib="NTA Mouse 15K cDNA Clone Set"  
/sex="Clones arrayed from a variety of cDNA libraries"  
/dev\_stage="Clones arrayed from a variety of cDNA libraries"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 160 a 169 c 168 g 130 t

## Alignment Scores:

Pred. No.: 5.81e-196 Length: 627  
Score: 199.00 Matches: 199  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.50% Indels: 0  
DB: 12 Gaps: 0

US-09-964-956-13 (1-1896) x BC086250 (1-627)

QY 1652 TrpHisLeuValArgHisGluHisGlyAspGlnLysGluGlyAspArgGlySerLys 1671  
|||||  
DB 31 TGGCATCTGGTGAAGAACCATGACATGAGACACAGAGAGGGGCGGGGACCAAG 90  
QY 1672 MetValSerGluLeuThrLeuThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPhe 1691  
|||||  
DB 91 ATGGTGCTGTGACATCTATCTGACACGACTATTGGCCACAAAGGGCACACTGCGAGAAGTTT 150  
QY 1692 ValAspAspLeuPheGluThrIlePheSerThrAlaHisArgGlySerAlaLeuProLeu 1711  
|||||  
DB 151 GTGGATGACCTATTTCAGACCATCTTCAGTACGGCCACCGTGGCTCCGCTCTGCCGCTA 210  
QY 1712 AlaIleLysTyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIleHisAsp 1731  
|||||  
DB 211 GCCATCAAGTACATGTTTGACTCTCTCGATGACGAGCTGATAAGCAGCGCATCCATGAC 270

QY 1732 ProHisValArgHisThrTrpLysSerAsnCysLeuProLeuArgPheTrpValAsnMet 1751  
|||||  
DB 271 CCGCAGCTCGCCACACCTGGAAGCAACTGCCTGCCCTACGGTTTGGGTGAATATG 330  
QY 1752 IleLysAsnProGlnPheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeu 1771  
|||||  
DB 331 ATCAAGAACCTCAGTTTGTGTGTGACATCCATAGAAGACAGCATCAGCGATGCTCTC 390  
QY 1772 SerValValAlaGlnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLys 1791  
|||||  
DB 391 TCTGTGGTGGCCAGACCTTCATGGACTCCCTGCTCCACTTCGGAGCACCGATGGCGAAG 450  
QY 1792 AspSerProSerAsnLysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpVal 1811  
|||||  
DB 451 GATTCCCTCCCAACAGCTCTCTATGCCAAGGACATCCCGACGACACAGAACTGGGTA 510  
QY 1812 GluArgTyrTyrSerAspIleGlyLysMetProIleSerAspGlnAspMetAsnAla 1831  
|||||  
DB 511 GAGAGGTATTATTTCAGACATTGGGAAGATGCGGCAATAAGTGACGACATGAATGCA 570  
QY 1832 TyrLeuAlaGluGlnSerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeu 1850  
|||||  
DB 571 TACTGGCCGAGCAGTCCGGATGCATGATGATTCAATACATGATGATGCGCTC 627

## RESULT 2

AL707761 483 bp mRNA linear EST 22-MAR-2002  
LOCUS DKFZp686H2349\_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
DEFINITION DKFZp686H2349 5', mRNA sequence.  
ACCESSION AL707761  
VERSION AL707761.1 GI:19691116  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 483)  
AUTHORS Wambutt,R., Heubner,D., Mewes,W., Well,B. and Wiemann,S.  
TITLE EST (Wambutt,R., Heubner,D., Mewes,W., Well,B. and Wiemann,S.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Wambutt R  
MIPS Am Klopferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.  
No sl sequence available.  
This clone (DKFZp686H2349) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

### Source

1. 483  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFZp686H2349"  
/clone\_lib="686 (synonym: hlcc3)"  
/tissue\_type="human skeletal muscle"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiI; B:  
cDNA-collection"  
BASE COUNT 109 a 147 c 142 g 85 t  
ORIGIN

## Alignment Scores:

Pred. No.: 3.69e-147 Length: 483  
Score: 152.00 Matches: 152  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.02% Indels: 0

DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x AL707761 (1-483)

QY 695 ValLysLeuProGluAspCysProGlnLeuLeuArgValAspLysIleLeuValProVal 714  
|||||  
Db 3 GTGAAGCTGCCGAGGACTGCCCCAGCTGCTCGAGTGGACAGATCCTGGTGGCCGTG 62  
QY 715 GluValIleLysProIleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGln 734  
|||||  
Db 63 GAGGTGATCAAGCTATCAGCTGAAGGCCAAGAACCTCCCCAGCCAGCTGGGCAG 122  
QY 735 ArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArg 754  
|||||  
Db 123 CGTGGCTAGCAATGCATCTCAACATTCAGGCGAGGAGCAGGAGTGCCTGGCTGCGC 182  
QY 755 PheAsnSerSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIle 774  
|||||  
Db 183 TTCAACAGCTCCAGCTACAGTCCCAAGACACCTCTATTCTTATGAAGGATGGAGATC 242  
QY 775 AsnAsnLeuProValGluLeuThrValTrpAsnGlyHisPheAsnIleAspAsnPro 794  
|||||  
Db 243 AACAACTGCCCGTGAGTTGACAGTCGTGTGGAATGGGCACATTCACATTTGACACCCA 302  
QY 795 AlaGlnAsnLysValHisLeuTyrLysCysGlyAlaMetArgGluSerCysGlyLeuCys 814  
|||||  
Db 303 GCTCAGAATAAGTTACCTCTACAAGTGTGGAGCCATGCTGAGAGCTCGGGCTGTGC 362  
QY 815 LeuLysAlaAspProAspPheAlaCysGlyTrpCysGlnGlyProGlyGlnCysThrLeu 834  
|||||  
Db 363 CTCAGCTGACCCAGACTTCGATGTGGTGTGGTGGCCAGGCCAGCCAGTGCACCTG 422  
QY 835 ArgGlnHisCysProAlaGlnGluSerGlnTrpLeu 846  
|||||  
Db 423 CGCCAGCACTGCCCTGCCCCAGGAGAGCCAGTGGCTT 458

RESULT 3  
BM105323  
LOCUS 567 bp mRNA linear EST 21-NOV-2001  
DEFINITION 508753 MARC 3BOV Bos taurus cdNA 5', mRNA sequence.  
ACCESSION BM105323  
VERSION BM105323.1 GI:17036393  
KEYWORDS EST.  
SOURCE cow.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 567)  
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett  
,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G.,  
Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cdNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
21180013  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCAGTCCAGCAGC  
Plate: 100 row: F column: 12  
Seq primer: ATTTAGTGCATATAG.  
Location/Qualifiers

DB: 1..567  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 3BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from marrow, alveolar  
macrophage, ovary, fetal semitendinosus muscle, and fetal  
longissimus muscle."  
BASE COUNT 139 a 188 c 134 g 106 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.71e-132 Length: 567  
Score: 138.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.28% Indels: 0  
DB: 13 Gaps: 0

US-09-964-956-13 (1-1896) x BM105323 (1-567)

QY 1757 PheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGln 1776  
|||||  
Db 1 TTCTGTTTGTGACATCCACAAGAACAGCATCACCGAGCTCTCCCTCGTGGTGGCCAG 60  
QY 1777 ThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsn 1796  
|||||  
Db 61 ACCTTCTATGGACTCGTGTCCACATCGGAGCAGCCGCTGGCAAGGACTCCCTCCCAAC 120  
QY 1797 LysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSer 1816  
|||||  
Db 121 AAGCTGCTCTATGCCAAGGACATCCCCACGACTACAGACACTGGTGGAGAGTATTACTCA 180  
QY 1817 AspIleGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluGln 1836  
|||||  
Db 181 GACATCGGGAAGATCGCCGCATCAGCGACACGAGCATGAACGCGTAGCTAGCTAGCAG 240  
QY 1837 SerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyr 1856  
|||||  
Db 241 TCCCGGATGCATATGATGAGTTCAACACCATGAGCGCCCTCTCCGAGATCTTCTCTCTAC 300  
QY 1857 ValGlyLysTyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCysGlyLys 1876  
|||||  
Db 301 GTGGCAAGTACAGTACAGTACAGATCCTCGCACCCCTGGACACCATGACCATGTTGGGAG 360  
QY 1877 GlnLysLeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerLeuAspSer 1894  
|||||  
Db 361 CAGAAACTGGCCCTACAACTAGAACAAAGTCATAACCCCTCATGAGCTTAGACAGC 414

RESULT 4  
BF952116/c  
LOCUS BF952116  
DEFINITION OVI-NN0229-101100-445-e08 NN0229 Homo sapiens cdNA, mRNA sequence.  
ACCESSION BF952116  
VERSION BF952116.1 GI:12369391  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 512)  
REFERENCE 1  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
TITLE  
JOURNAL  
MEDLINE  
FEATURES

COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1&tl2=QV1-NN0229-  
101100-445-e08&tl3=2000-11-10&tl4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 18  
High quality sequence stop: 512.

FEATURES  
source  
1..512  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN0229"  
/dev\_stage="Adult"  
/note="Organ: nervous\_normal; Vector: puc18; Site.1: SmaI;  
Site.2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 96 a 132 c 147 g 137 t  
ORIGIN

Alignment Scores:  
Pred. No.: 8.35e-114 Length: 512  
Score: 120.00 Matches: 120  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.33% Indels: 0  
DB: 12 Gaps: 0

US-09-964-956-13 (1-1896) x BF952116 (1-512)

QY 1775 AlaGlnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerPro 1794  
|||||  
Db 495 GCTCAGACTTCATGGACTCTGTCTCCACGTACAGACCCGCTGGGCAAGAGCTCGCCC 436  
|||||

QY 1795 SerAsnLysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyr 1814  
|||||  
Db 435 TCCAAACAGCTGCTGTATGCCAAGGACATCCCGACGTACAAGATTGGTGGAGAGGTAT 376  
|||||

QY 1815 TyrSerAspIleGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAla 1834  
|||||  
Db 375 TACTCAGACATAGGGAAGATCCAGCCATCAGCCACCAAGACATGAACGCATACCTGGCT 316  
|||||

QY 1835 GluGlnSerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePhe 1854  
|||||  
Db 315 GAGCAGTCCCGGATGCATGATGAATGATTCACACCATGAGTGCCTCTCAGAGATCTTC 256  
|||||

QY 1855 SerTyrValGlyLysTyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCys 1874  
|||||  
Db 255 TCCTATGTGGCAATACAGCGAGGAGATCCTTGGACCTCTGGACACGACGACGAGTGT 196  
|||||

QY 1875 GlyLysGlnLysLeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerLeuAspSer 1894  
|||||  
Db 195 GGGAGGAGAGAACTGGCGCTACAACTAGAAAGTCTAATACCTCATGAGCTTAGACAGC 136  
|||||

RESULT 5  
BF953019  
LOCUS BF953019 564 bp mRNA linear EST 22-JAN-2001  
DEFINITION QV3-NN0198-111100-373-h10 NN0198.Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF953019  
VERSION BF953019.1 GI:12370294  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 564)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&tl2=QV3-NN0198-  
111100-373-h10&tl3=2000-11-11&tl4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 19  
High quality sequence stop: 556.  
High quality sequence stop: 556.

FEATURES  
Location/Qualifiers  
1..564  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN0198"  
/dev\_stage="Adult"  
/note="Organ: nervous\_normal; Vector: puc18; Site.1: SmaI;  
Site.2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 114 a 171 c 176 g 103 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.68e-102 Length: 564  
Score: 109.00 Matches: 109  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.75% Indels: 0  
DB: 12 Gaps: 0

US-09-964-956-13 (1-1896) x BF953019 (1-564)

QY 1 MetLysAlaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMet 20  
|||||  
Db 234 ATCAAAGCCATGCCCTGGAACTGGACCTGCCTCTCTCCACCTCTCATGGTGGCATG 293  
|||||

QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
|||||  
Db 294 GCGTCTCTCCACTTGTCTACCCGCGCAGCCCGCTGTCCAGAGACGAGCGTCAATT 353  
|||||

QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
|||||  
Db 354 GTCACATTCCGAGGAGAGCCCGCGAGGGTTTCAATCACCTGGTGTGATGAGAGACA 413  
|||||

QY 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
|||||  
Db 414 GGACACATTACTTGGGGCCCGTCAATCGGATTTACAACTCTCCACGCGACCTGAAGTC 473  
|||||

QY 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100  
|||||

Db 474 TTGGTGACGATGAGACAGCGCGGAGGAGAGACAAACCCCAAGTGTATCCCAACCCCGCATC 533

QY 101 ValGlnThrCysAsnGluProLeuThr 109  
 |||||  
 Db 534 GTCAGACCTGCAATGAGCCCTGACC 560

## RESULT 6

BB656076 612 bp mRNA linear EST 26-OCT-2001  
 LOCUS BB656076 RIKEN full-length enriched, 12 days embryo spinal ganglion  
 DEFINITION Mus musculus cDNA clone DI30060F15 5', mRNA sequence.  
 ACCESSION BB656076  
 VERSION BB656076.1 GI:16489904  
 KEYWORDS EST.  
 SOURCE house mouse.

## ORGANISM

Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 612)  
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,  
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,  
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,  
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,  
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,  
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,  
 Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
 Unpublished (2001)

## TITLE

JOURNAL RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,  
 M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10): 1617-1630 (2000)  
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
 Wataniki,M., Yoneda,Y., Shikawa,T., Ozawa,K., Tanaka,T., Matsuura,  
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
 Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,  
 Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,  
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and  
 Hayashizaki,Y.  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.

## FEATURES

Location/Qualifiers  
 1..612  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="DI30060F15"  
 /clone\_lib="RIKEN full-length enriched, 12 days embryo  
 spinal ganglion"  
 /tissue\_type="spinal ganglion"  
 /dev\_stage="12 days embryo"  
 /lab\_host="DH10B"  
 /note="Site\_1: Sali; Site\_2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGCGCGCAACTCAGTGTCTTTTCTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. Second strand cDNA was prepared with the  
 primer adapter of sequence [5'  
 GAGAGAGATTCGAGTGAATTAATTAATCCCCCCCCCC 3']. cDNA  
 was cleaved with BamHI and XhoI. Vector: a modified  
 pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 147 a 176 c 160 g 129 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,34e-98 Length: 612  
 Score: 105.00 Matches: 203  
 Percent Similarity: 99.02% Conservative: 0  
 Best Local Similarity: 99.02% Mismatches: 1  
 Query Match: 5.54% Indels: 2  
 DB: 10 Gaps: 0

US-09-964-956-13 (1-1896) x BB656076 (1-612)

QY 1043 ValAtGileGluProGluTrpSerIleValSerGlyAsnThrProIleAlaValTrpGly 1062  
 |||||  
 Db 1 GTGCGGATTGAACCAAGATGGAGCATGTTCAGTGGGAACACACCTATTGTCTGGGGA 60  
 QY 1063 ThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGlyLysGluHis 1082  
 |||||  
 Db 61 ACTCACCTGGACCTCATACAGAACCCACACATCCGTGCCAAGCATGGAGGAAGACAC 120  
 QY 1083 IleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAlaProAlaLeuAla 1102  
 |||||  
 Db 121 ATCAACATCTGTGAGGTCTCTAAATGCTACAGATGACCTGCCAGGCTCCAGCCCTTGC 180  
 QY 1103 LeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPheGlyPheIleLeu 1122  
 |||||  
 Db 181 CTGGGTCCCGACCCAGTCAGATCATCTGAGAGGCTGAAGAAATTTGGTTTCATCCTG 240  
 QY 1123 AspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsnPhetThrTyrProAsnPro 1142  
 |||||  
 Db 241 GACAATGTCCAGTCTCTGCTAATTTCTCAACAAGACCAACTTCACCTATTATCTTAACCC 300  
 QY 1143 ValPheGluAlaPheGly-ProSerGlyIleLeuGluLeuLysProGlyThrProIleIle 1162  
 |||||  
 Db 301 GTGTTTGGAGCTTTCAG-TCCCTCGGGAATCCTGGAGCTCAAGCGGGCACCCCATAT 359  
 QY 1162 eLeuLysGlyLysAsnLeuIleProValAlaGlyGlyAsnValLysLeuAsnTyrThr 1182  
 |||||  
 Db 360 CCTAAGGCAAGAACTTGATCCCGCTGTGGCGGAGGCAATGTGAAGTGAACTACAC 419  
 QY 1182 rValLeuValGlyGluLysProCysThrValThrValSerAspValGlnLeuLeuCysGI 1202  
 |||||  
 Db 420 CGTGTGTTGGGAGAAACCATGCACCGTGACAGTATCAGATGTGCAGCTGCTCGCA 479  
 QY 1202 uSerProAsnLeuIleGlyArgHisLysValMetAlaArgValGlyGlyMetGluTyrSe 1222  
 |||||  
 Db 480 GTCTCCCAACCTCATCGGAGGCACAAAGTGAATGGCACGGGTGGGTGGCATGATTC 539  
 QY 1222 rProGlyMetValTyrIleAlaProAspSerProLeuSerLeuProAlaIleValSerIl 1242  
 |||||  
 Db 540 CCTCGGATGGTATATACATGCCCCACAGACACCCGCTCAGCTGCCCGCCCATTTAGCAT 599  
 QY 1242 eAlaValAlaGly 1246  
 |||||  
 Db 500 CGCGGTGGCTGGG 612  
 RESULT 7  
 BG694413



QY 590 GlyValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValGlyAsnGln 609  
 Db 230 GGCGTCACCTGACCTTTTCAGGACCTTCAGAGATGGATGGCTGGTCTGGGCAATCAG 171  
 QY 610 IleGlnCysTyrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAsp 629  
 Db 170 ATCCAGTCTACTCCCTCGAGCCAGAGGTGCCCGGATCATACAGAGATGGGAC 111  
 QY 630 HisHisValValGlnLeuGlnLeuLysSerTysGluThrGlyMetThrPheAlaSerThr 649  
 Db 110 CACCATGTGCTACAGCTTCAGCTCAATCAAAAGGAGACGGCGCATGACCTTCGCCAGCACC 51  
 QY 650 SerPheValPheTyrAsnCysSerValHisAsnSerCysLeuSerCys 665  
 Db 50 AGCTTTGTCTTCTACAATTGCAGCGTCCACAATTCGTGCTGCTGCTGC 3

RESULT 9  
 AW018699  
 LOCUS  
 DEFINITION f049b01.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone  
 IMAGE:3733033 5' similar to TR:Q14988 Q14988 OCT PROTEIN ;, mRNA  
 sequence.  
 ACCESSION AW018699  
 VERSION AW018699.1 GI:5872229  
 KEYWORDS EST.  
 SOURCE zebrafish.  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
 ; Cyprinidae; Danio.  
 1 (bases 1 to 548)  
 Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy  
 ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood  
 ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,  
 Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,  
 Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.  
 and Wilson,R.  
 WashU Zebrafish EST Project 1998  
 Unpublished (1998)  
 Other\_ESTs: f049b01.x1  
 Contact: Stephen L. Johnson  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: zbrafish@watson.wustl.edu  
 cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:  
 Matthew Clark. DNA Sequencing by: Washington University Genome  
 Sequencing Center Clone distribution: Genome Systems, St. Louis,  
 Missouri (web address: www.genomesystems.com) (email contact:  
 info@genomesystems.com) and Research Genetics, Huntsville, Alabama  
 (web address: www.resgen.com) (email contact: info@resgen.com) and  
 RessourcenzentrumPrimarDatenbank, Berlin, Germany (web address:  
 www.rzpd.de)  
 Seq primer: T3 ET from Amersham  
 High quality sequence stop: 498.  
 Location/Qualifiers  
 1..548  
 /organism="Danio rerio"  
 /db\_xref="taxon:7955"  
 /clone="IMAGE:3733033"  
 /clone\_lib="Zebrafish WashU MPIMG EST"  
 /sex="mixed"  
 /tissue\_type="26 somite embryos, adult livers, shield  
 stage embryos"  
 /lab\_host="xLi-blue MRF"  
 /note="Vector: pSPOR1; Site\_1: NotI; Site\_2: SalI; 1st  
 strand cDNA was primed with a Not I - oligo(gr)15 primer  
 [5'-pCAGTCTAGATCGGAGCGCGCCCTTTTCTTTTCTTTT3'];  
 double-stranded cDNA was ligated to Sal I adaptors (BRL),  
 digested with Not I and cloned into the Not I and Sal I  
 sites of the pSPOR1 vector (BRL). Library was constructed  
 by Matthew Clark (Lehrach lab; ICRF, London and Max Planck

Institut fuer Molekulare Genetik, Berlin). cDNAs for EST  
 analysis were selected following oligonucleotide  
 hybridization fingerprinting of arrayed clones from  
 zebrafish late somitogenesis (26 ss), adult liver or  
 embryonic shield stage (5-6 h) libraries. Fingerprint  
 data were used to computationally cluster cDNAs, and a  
 single cDNA from each cluster was chosen for sequencing.  
 In some cases multiple members of the same cluster were  
 sequenced to assess clustering parameters or single clones  
 were sequenced additional times to assess quality  
 control."

BASE COUNT 129 a 140 c 144 g 135 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 6.58e-67 Length: 548  
 Score: 75.00 Matches: 75  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.96% Indels: 0  
 DB: 1.0 Gaps: 0

US-09-964-956-13 (1-1896) x AW018699 (1-548)

QY 1256 ValLeuIleAlaTyrLysArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGln 1275  
 Db 138 GTCTTAATTCCTACAAGCGCAAGTCCGTCGAGAGTACCTCACCTCAAGCGCTGCAG 197  
 QY 1276 MetClnMetAspAsnLeuGluSerArgValAlaLeuGluCysLysGluAlaPheAlaGlu 1295  
 Db 198 ATCAGATGACGACCACTGGAGTCTAGAGTAGCCCTAGAGTAAAGAGGCATTCCTCGAG 257  
 QY 1296 LeuGlnThrAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeu 1315  
 Db 258 CTCACACGCGCATTCATGACTGCACAGTATCTGGATGGAGCTGGTATACCATTTCTG 317  
 QY 1316 AspTyrArgThrTyrThrMetArgValLeuPheProGlyIleGlu 1330  
 Db 318 GACTATCGCACATATACATACGCGGTCTCTGTTCTTCTGGAATCGAG 362

RESULT 10

AA077556 233 bp mRNA linear EST 24-SEP-1999  
 LOCUS 7B19F10 Chromosome 7 Fetal Brain cDNA Library Homo sapiens cDNA  
 DEFINITION clone 7B19F10, mRNA sequence.

ACCESSION AA077556  
 VERSION AA077556.1 GI:1837030  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 233)  
 Touchman,J.W., Bouffard,G.G., Weintraub,L.A., Idol,J.R., Wang,L.,  
 Robbins,C.M., Nussbaum,J.C., Lovett,M. and Green,E.D.

2006 expressed-sequence tags derived from human chromosome

7-enriched cDNA libraries

Genome Res. 7 (3), 281-292 (1997)

JOURNAL 9728905

MEDLINE

COMMENT

Contact: Eric D. Green

Genome Technology Branch

National Human Genome Research Institute/NIH

49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892

Tel: 3014020201

Fax: 3014024735

Email: egreen@nhgri.nih.gov

Plate: 19 row: F column: 10

Seq primer: -21M13 (ABI).

Location/Qualifiers

1..233

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="7B19F10"

FEATURES  
 source

```

/clone_lib="Chromosome 7 Fetal Brain cDNA Library"
/sex="female and male mixture"
/tissue_type="brain"
/dev_stage="pool of 9 week and 12 week"
/lab_host="E. coli strain DH5 alpha"
/notes="Organ: brain; Vector: pAMP10; cDNA was generated
from cytoplasmic RNA using a mixture of random DNA
hexamers and oligo(dT). From this pool of cDNA, human
chromosome 7-enriched cDNA was isolated by direct cDNA
selection using chromosome 7 genomic DNA (cosmids). The
resulting direct-selected cDNA was cloned into a plasmid
vector using a non-directional uracil DNA glycosylase (UDG
)-mediated cloning strategy."
BASE COUNT      50 a      63 c      75 g      45 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      3,35e-63      Length:      233
Score:          71.00      Matches:      71
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      3,74%      Indels:      0
DB:              9      Gaps:      0

```

US-09-964-956-13 (1-1896) x AA077556 (1-233)

```

QY 1311 GlyileProPheLeuAspTyrArgThrTyrMetArgValLeuPheProGlyIleGlu 1330
DB 231 GGGATTCGGTCTCGACTATAGACTTACACATGCGGGTCTGTTCACGGAATGAA 172
QY 1331 AspHisProValLeuArgAspLeuGluValProGlyTyrArgGlnGluArgValGluLys 1350
DB 171 GACCACCTGTCTCCGGGACCTTGAGTGTCGGGCTACCGCAGGAGCGTGTGGAGAAA 112
QY 1351 GlyLeuLysLeuPheAlaGlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArg 1370
DB 111 GGCCTGAAGCTTCCGCCAGCTCATCAACAAGGTGTCTGTCTTCATCCGC 52
QY 1371 ThrLeuGluSerGlnArgSerPheSerMetArg 1381
DB 51 AGCCTTGAGTCCAGCGTAGCTTCTCCATGCGC 19

```

```

RESULT 11
AA077578/c
LOCUS      7B19H10 Chromosome 7 Fetal Brain cDNA Library Homo sapiens cDNA
DEFINITION clone 7B19H10, mRNA sequence.
ACCESSION  AA077578
VERSION     AA077578.1  GI:1837052
KEYWORDS   EST.
SOURCE     human.

```

```

ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 231)
AUTHORS   Touchman,J.W., Bouffard,G.G., Weintraub,L.A., Idol,J.R., Wang,L.,
            Robbins,C.M., Nussbaum,J.C., Lovett,M. and Green,E.D.
            2006 expressed-sequence tags derived from human chromosome
            7-enriched cDNA libraries
TITLE      Genome Res. 7 (3), 281-292 (1997)
JOURNAL    97228905
MEDLINE    Contact: Eric D. Green
            Genome Technology Branch
            National Human Genome Research Institute/NIH
            49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
            Tel: 3014020201
            Fax: 3014024735
            Email: egreen@nhgri.nih.gov
            Plate: 19 row: H column: 10
            Seq primer: -21M13 (ABI).
            Location/Qualifiers
FEATURES   1..251
            /organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/clone="7B19H10"
/clone_lib="Chromosome 7 Fetal Brain cDNA Library"
/sex="female and male mixture"
/tissue_type="brain"
/dev_stage="pool of 9 week and 12 week"
/lab_host="E. coli strain DH5 alpha"
/notes="Organ: brain; Vector: pAMP10; cDNA was generated
from cytoplasmic RNA using a mixture of random DNA
hexamers and oligo(dT). From this pool of cDNA, human
chromosome 7-enriched cDNA was isolated by direct cDNA
selection using chromosome 7 genomic DNA (cosmids). The
resulting direct-selected cDNA was cloned into a plasmid
vector using a non-directional uracil DNA glycosylase (UDG
)-mediated cloning strategy."
BASE COUNT      52 a      75 c      74 g      50 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      4,89e-60      Length:      251
Score:          68.00      Matches:      68
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      3,59%      Indels:      0
DB:              9      Gaps:      0

```

US-09-964-956-13 (1-1896) x AA077578 (1-251)

```

QY 548 ProArgArgPheAlaSerGluMetLysGlnCysValArgLeuThrValHisProAsnAsn 567
DB 206 CCCCGCAGGTTGCTCGAGATGAAGCAGTGTGTCCGGCTGACGGTCCATCCCAACAT 147
QY 568 IleSerValSerGlnTyrAsnValLeuValLeuGluThrTyrAsnValProGluLeu 587
DB 146 ATCTCCCTCTCTAGTACACGCTGCTGTCTGTGGAGAGCTACAATGTCCTCCGAGCTG 87
QY 588 SerAlaGlyValAlaAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValGly 607
DB 86 TCAGCTGGCGTCAACTGCACCTTTGAGGACCTGTGAGAGATGGATGGCTGCTGCTGGGC 27
QY 608 AsnGlnIleGlnCysTyrSerPro 615
DB 26 AATCAGATCCAGTCTACTCCCT 3

```

RESULT 12

```

BF953008/c
LOCUS      BF953008      493 bp      mRNA      linear      EST 22-JAN-2001
DEFINITION QV3-NN0198-111100-373-b07 NN0198 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF953008
VERSION     BF953008.1  GI:12370283
KEYWORDS   EST.
SOURCE     human.

```

```

ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 493)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,C.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
TITLE      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL    20202663
MEDLINE    Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001

```

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-NN0198-111100-373-b07&t3=2000-11-11&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 493.  
Location/Qualifiers  
1. .493

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN0198"  
/dev\_stage="Adult"  
/note="Organ: nervous\_normal; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 88 a 156 c 150 g 99 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.36e-57 Length: 493  
Score: 66.00 Matches: 80  
Percent Similarity: 98.77% Conservative: 0  
Best Local Similarity: 98.77% Mismatches: 1  
Query Match: 3.48% Indels: 1  
DB: 12 Gaps: 0

US-09-964-956-13 (1-1896) x BF9533008 (1-493)

QY 1 MetLysAlaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMet 20  
|||||  
Db 278 ATGAAGCCATGCCCTGGAACTGGACCTGCTTCTCTCCACCTCCCTCATGTGGGCATG 219  
QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
|||||  
Db 218 GGTCTCTCCACTTTGTCTACCCGGCAGCCCGCTGCTCCAGAAAGCAGCGGTCAATTT 159  
QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
|||||  
Db 158 GTCAATCTCCAGAGAGAGCCCGGAGGGTTTCAATCACCCTGGTGGATGAGAGGACA 99  
QY 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
|||||  
Db 98 GGACACATTTA-CTTGGGGCGTCAATCGGATTTACAAAGCTCTCCAGCGACCTGAAGTC 40  
QY 81 Leu 81  
|||  
Db 39 TTG 37

## RESULT 13

AL595528  
LOCUS BQ434703  
DEFINITION AGENCOURT\_7909800 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6156354, mRNA sequence.

ACCESSION AL595528

VERSION AL595528.1

GI:15007603

KEYWORDS EST.

SOURCE western clawed frog.

ORGANISM Silurana tropicalis

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Silurana.

1 (bases 1 to 650)

Huckle, E., Taylor, R., McMurray, A., Ashurst, J.L., Zorn, A.M. and

Rogers, J.

Sanger Xenopus tropicalis EST project 2001

Unpublished (2001)

Contact: Huckle E

Sanger Centre

Hinton, Cambridgeshire, CB10 1SA, UK  
Email: trop@sanger.ac.uk  
Sanger Xenopus tropicalis EST project 2001  
TROPICALIS\_SEQUENCE\_ID: Tgas005n23.sp6  
Sequencing primer: SP6  
This sequence is from a Xenopus Gene Collection (XGC) library  
constructed by Aaron M. Zorn.  
Location/Qualifiers  
1. .650

FEATURES  
source

/organism="Silurana tropicalis"  
/db\_xref="taxon:8364"  
/clone\_lib="Tgas005n23"  
/dev\_stage="XGC-gastrula"  
/dev\_stage="gastrula (stages 10.5-13 mixed)"  
/lab\_host="Escherichia coli XL1-blue"  
/note="Vector: PCS107; Site\_1: EcoRI; Site\_2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into PCS107 with EcoRI at the 5' end and NotI at the 3' end."  
BASE COUNT 180 a 174 c 161 g 135 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.55e-54 Length: 650  
Score: 63.00 Matches: 63  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.32% Indels: 0  
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x AL595528 (1-650)

QY 1665 GluGlyAspArgGlySerLysMetValSerGluIleTyrLeuThrArgLeuAlaThr 1684  
|||||  
Db 332 GAGGGCGATAGAGGTAGCAAGATGGTCTGTGAGATCTACTGCTGCTTGGCTTACA 391  
QY 1685 LysGlyThrLeuGlnLysPheValAspLeuPheGluThrIlePheSerThrAlaHis 1704  
|||||  
Db 392 AAGGGAACCTCTCAGAAAGTTTGTGGATGATCTCTTTGAGACAATCTTTCAGCAGCGCAC 451  
QY 1705 ArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuAspGluAla 1724  
|||||  
Db 452 CTTGGAAGCGCGCTCCCGCTCGCTATTAAATACATGTTGATTCTCTGGCAGCAGCGG 511  
QY 1725 AspLysHis 1727  
|||||  
Db 512 GACAGCAT 520

## RESULT 14

BQ434703

LOCUS

DEFINITION

AGENCOURT\_7909800 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6156354, 5' mRNA sequence.

ACCESSION BQ434703

VERSION BQ434703.1

GI:21173779

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 844)

NIH-MGC http://imgc.ncbi.nlm.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA library preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
Plate: L1AM13500 row: h column: 19  
High quality sequence stop: 597.

# FEATURES

source  
1. .844  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6156354"  
/lab\_host="NIH\_MGC\_71"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2.1 kb."  
BASE COUNT 202 a 256 c 246 g 140 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.51e-54 Length: 844  
Score: 63.00 Matches: 63  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.32% Indels: 0  
DB: 14 Gaps: 0

US-09-964-956-13 (1-1896) x BQ434703 (1-844)

QY 1665 GluGlyAspArgGlySerLysMetValSerGluIleTyrLeuThrArgLeuLeuAlaThr 1684  
|||||  
DB 55 GAGGTGACCGCGCAGACAGATGGTCTCGGAGATCTACTTGACACGGCTACTGGCCACC 114  
QY 1685 LysGlyThrLeuGlnLysPheValAspLeuPheGluThrIlePheSerThrAlaHis 1704  
|||||  
DB 115 AAGGGCACACTGCAGAGTTTGTGGACGACCTGTTGAGACCACTTTCAGACGCGCACAC 174  
QY 1705 ArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAla 1724  
|||||  
DB 175 CGGGGCTCAGCCCTGCGCGCTGGCCATCAAGTACATGTTCTGACCTCTCTGGATGAGCAGGCC 234  
QY 1725 AspLysHis 1727  
|||||  
DB 235 GACAGGCAC 243

# RESULT 15

BQ878090  
LOCUS BQ878090 851 bp mRNA linear EST 16-AUG-2002  
DEFINITION AGENCOURT 8072266 NIH\_MGC\_112 Homo sapiens cDNA clone IMAGE:6089613  
5', mRNA sequence.

ACCESSION BQ878090  
VERSION BQ878090.1 GI:22270098  
KEYWORDS EST.  
SOURCE human.

# ORGANISM

Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 851)

# REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: L1CM2329 row: k column: 22  
High quality sequence stop: 625.

# FEATURES

source  
1. .851  
Location/Qualifiers  
/organism="Homo sapiens"

/db\_xref="taxon:9606"  
/clone="IMAGE:6089613"  
/lab\_host="NIH\_MGC\_112"  
/tissue\_type="melanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."  
BASE COUNT 189 a 288 c 236 g 138 t  
ORIGIN

# Alignment Scores:

Pred. No.: 3.55e-54 Length: 851  
Score: 63.00 Matches: 63  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.32% Indels: 0  
DB: 14 Gaps: 0

US-09-964-956-13 (1-1896) x BQ878090 (1-851)

QY 1665 GluGlyAspArgGlySerLysMetValSerGluIleTyrLeuThrArgLeuLeuAlaThr 1684  
|||||  
DB 430 GAGGTGACCGCGCAGCAGATGGTCTCGGAGATCTACTTGACACGGCTACTGGCCACC 489  
QY 1685 LysGlyThrLeuGlnLysPheValAspLeuPheGluThrIlePheSerThrAlaHis 1704  
|||||  
DB 490 AAGGGCACACTGCAGAGTTTGTGGACGACCTCTTTGAGACCACTTTCAGCAGCGCACAC 549  
QY 1705 ArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAla 1724  
|||||  
DB 550 CGGGGCTCAGCCCTGCGCGCTGGCCATCAAGTACATGTTCTGACCTCTCTGGATGAGCAGGCC 609  
QY 1725 AspLysHis 1727  
|||||  
DB 610 GACAGGCAC 618

Search completed: June 28, 2003, 23:58:02  
Job time : 5367 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 28, 2003, 04:02:59 ; Search time 10687 Seconds  
(without alignments)  
5163.181 Million cell updates/sec

Title: US-09-964-956-13  
Perfect score: 9990  
Sequence: 1 MKAMPWNWTCLLSHLLVMG.....QKLAYKLEQVITIMSLDSNK 1896

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09964956/runat\_24062003.124338.5205/app\_query\_fasta\_1.2055  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09964956.ecgn\_1.1.7475.ecgnat\_24062003.124338.5205 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEF\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	6794	68.0	6387	6	AX405895	AX405895 Sequence
2	6790	68.0	6263	9	AB007932	AB007932 Homo sapi
3	6765	67.7	6730	6	AX305728	AX305728 Sequence
4	6765	67.7	6730	10	D86949	D86949 Mus musculu
5	6393.5	64.0	6163	5	XELPLEX	D38175 Xenopus lae
6	6337	63.4	6466	10	D86948	D86948 Mus musculu
7	5991	60.0	6039	9	HSSEXGENE	X87852 H.sapiens m
8	5984.5	59.9	5982	10	D86950	D86950 Mus musculu
9	5847.5	58.5	5263	9	HSNOVPROT	X87832 Homo sapien
10	4986	49.9	3985	9	HSOCTPROT	X87831 Homo sapien
11	4421	44.3	198653	2	AC091474	AC091474 Mus muscu
12	4403	44.1	204152	2	AL807376	AL807376 Mus muscu
13	4394	44.0	219447	9	HUMFLNG6PD	L41410 Homo sapien
14	4378	43.8	160656	2	AL592072	AL592072 Homo sapi
15	4372.5	43.8	193961	2	AC025784	AC025784 Mus muscu
16	4355.5	43.6	200322	2	AL808145	AL808145 Mus muscu
17	3953	39.6	5986	9	AB046770	AB046770 Homo sapi
18	3631	36.3	6602	3	AF106932	AF106932 Drosophill
19	3366.5	33.9	2884	9	BC009343	BC009343 Homo sapi
20	2994	30.0	156723	3	AC010838	AC010838 Drosophill
21	2973.5	29.8	214660	2	AC094668	AC094668 Rattus no
22	2932.5	29.4	6665	3	AF106933	AF106933 Drosophill
23	2869.5	28.7	21595	2	AC014617	AC014617 Drosophill
24	2869.5	28.7	298469	3	AE003846	AE003846 Drosophill
25	2858	28.6	6612	3	AB080022	AB080022 Caenorhab
26	2792.5	28.0	5859	9	BSA011415	AJ011415 Homo sapi
27	2710	27.1	7308	9	AB007867	AB007867 Homo sapi
28	2707	27.1	6408	6	AX088118	AX088118 Sequence
29	2707	27.1	6643	9	HSRNASAP	X87904 Homo sapien
30	2621.5	26.2	137841	3	AC010915	AC010915 Drosophill
31	2584	25.9	6145	9	AB033032	AB033032 Homo sapi
32	2584	25.9	6145	9	AF149019	AF149019 Homo sapi
33	2580	25.8	6252	6	AX088110	AX088110 Sequence
34	2580	25.8	6252	6	AX210612	AX210612 Sequence
35	2580	25.8	6252	6	AX335107	AX335107 Sequence
36	2580	25.8	6252	9	AB002313	AB002313 Human mRN
37	2475.5	24.8	6754	9	AB014520	AB014520 Homo sapi
38	2463.5	24.7	5892	6	AX088114	AX088114 Sequence
39	2447	24.5	3016	9	BC007934	BC007934 Homo sapi
40	2425.5	24.3	188873	2	AC073658	AC073658 Mus muscu
41	2409.5	24.1	2020	9	BC028744	BC028744 Homo sapi
42	2228	22.3	5310	10	BC007481	BC007481 Mus muscu
43	2084	20.9	179145	9	AC009785	AC009785 Homo sapi
44	2084	20.9	188090	9	AC011625	AC011625 Homo sapi
45	2071	20.7	179592	2	AC009723	AC009723 Homo sapi

ALIGNMENTS

RESULT 1



Db 1353 CCGGGCCATCACTTCGAGATCAAGAGCCCTCGAGTCTCTGATACAGCGCGAGGCCAA 1412  
QY 380 rLeuAspLeuAlaTrpLeuLysValLysAspIleProCysSerSerAlaLeuLeuThrII 400  
Db 1413 CTTGGAGCTCACTGCTGCTGGGAGGAGCTCCAGTGCACCAAGCGCTGTCCCAT 1472  
QY 400 eAspAspAsnPhcCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValAr 420  
Db 1473 CGATGATACTTCTGTGGAGTGACATCAACCAAGCCCTCGGAGCTCAACTCCAGTGA 1532  
QY 420 gGlyIleProValPheThrGluAspArgAspArgMetThrSerValIleAlaIleValTy 440  
Db 1533 GGCCTGACCTGTACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1592  
QY 440 rLysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLysLysIleArgVa 460  
Db 1593 CAACGCTACAGGCTGTTTCTGGGAGTAAAGTGGGCAAGTGAAGATTCGGGC 1652  
QY 460 lAspGlyProArgGlyAsnAlaLeuGlnThrGluThrValGlnValVal---AspProGl 479  
Db 1653 CGAGGTCCCGCCCATGCTGGGTCCAGTACGAGATGGTCTCTGTCTCAAGGACGGAAG 1712  
QY 479 yProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuThrIleMetSerGl 499  
Db 1713 CCGCATCTCCGGAGATGCGCTTCCATGATCAGCGTACTGTACGTCAATGCTGA 1772  
QY 499 uArgGlnLeuThrArgValProValGluSerCysGlyGlnThrGlnSerCysGlyGluCy 519  
Db 1773 GAGACAGGTCAACAGGCTCCCGTGGAGTATGTGAGCAGTATACGACTTGTGGGAGTG 1832  
QY 519 sLeuGlySerGlyAspProHisCysGlyTrpCysValLeuHisAsnThrCysThrArgLy 539  
Db 1833 CTTGAGCTCTGGGAGCCCTCACTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1892  
QY 539 sGluArgCysGluArgSerLysGluProArgPheAlaSerGluMetLysGlnCysVa 559  
Db 1893 GGACAAATGCCACAGCCCTGGGAACCTAATCGATTGCTGCGCATCAGCAGCTAGTGT 1952  
QY 559 lArgLeuThrValHisProAsnAsnIleSerValSerGlnThrAsnValLeuValLe 579  
Db 1953 GAGCCTTGAGTGCATCCAGCAGCAGTCTCAGTATCTGAGCAGCAGCGGTTGCTTAGCCT 2012  
QY 579 uGluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSe 599  
Db 2013 GGTAGTGAGTGCCTGATCTATCTGCGGTATGCGCTGTGCTGTGCTGTGCTGTGCTGTG 2072  
QY 599 rGluMetAspGlyLeuValValGlyAsnGlnIleGlnCysTyrSerProAlaAlaLysGl 619  
Db 2073 AGAGGTGGAGGCGAGGTGTCGGGAGGCGAGGTCATCTGATCTCAGCTGGGCGCAAGGA 2132  
QY 619 uValProArgIleIleThrGluAsnGlyAspHisHisValValcInLeuGlnLeuLysSe 639  
Db 2133 TGTCCCT---GTATCCCGCTGATCAAGACTGGTGTGGCTGGAGTACAGCTGAGGTC 2189  
QY 639 rLysGluThrGlyMetThrPheAlaSerThrSerPheValPheThrAsnCysSerValHI 659  
Db 2190 CAAGGAGACAGGAAGATATTCTGACACCGAGTCAAGTTTACAACTGCGAGTGCCTCA 2249  
QY 659 sAsnSerCysLeuSerCysValGluSerProTyrArgCysHisTrpCysLysTyrArgHI 679  
Db 2250 CCAACTGCTGCTCTGTCACAGCGCTTCCTCGCTGCCATGCTGGTGGCAAGTACCCAA 2309  
QY 679 sValCysThrHisAspProLysThrCysSerPheGlnGluGlyArgValLysLeuProGl 699  
Db 2310 CTTCTGCACTCATGAGCCCAACCACTGCTCTCTCCAGGAGGCGCGGATCAATATTTCAGA 2369  
QY 699 uAspCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysPr 719  
Db 2370 GGACTGTCCCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2429  
QY 719 oIleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCy 739  
Db 2430 AATCACCTTAAAGCGCGGAATCTGCCCAAGCGCAGCTCCGGCCAGGAGGAGGAGGAGG 2489

QY 739 sIleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArgPheAsnSerSerSe 759  
Db 2490 TGTCTCAACATACAAGGAGGAGCCTCCCGCTCTCGCTTCAACAGCTCCAG 2549  
QY 759 rValcInCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsnLeuProVa 779  
Db 2550 CGTTCACTGTGAGAACAGCTCGTACAGTATGATGATGATGATGATGATGATGATGATGAT 2609  
QY 779 lGluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsnProAlaGlnAsnLysVa 799  
Db 2610 GGATTTCTGCTGGGTGGAGCGCAATTTTCATCATTTGACACCTCTCAGAGCTGAAAGT 2669  
QY 799 lHisLeuTyrLysCysGlyAlaMetArgGluSerCysGlyLeuCysLeuLysAlaAspPr 819  
Db 2670 CCATCTCTACAAGTGTGAGCGCCACGCGGAGAGCTGGGCTCTGCCTCAAGGCCACCG 2729  
QY 819 oAspPheAlaCysGlyTrpCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysPr 839  
Db 2730 GAAGTTTCAAGTGTGGCTGTGTCAGCGCGGAGCGAGCGAGTGCACCTCCACGACCTGTAC 2789  
QY 839 oAlaGlnGluSerGlnTrpLeuGluLeuSerGlyAlaLysSerLysCysThrAsnProAr 859  
Db 2790 CAGCCCTTCCAGCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2849  
QY 859 gIleThrGluIleLeuProValThrGlyProArgGluGlyGlyThrLysValThrIleAr 879  
Db 2850 AATCACCAGATTTTACCGTGTCTGCGAGTCTCCGAGATCGCCACCATGTCAGTGTCAACCTCA 2909  
QY 879 gGlyGluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyVa 899  
Db 2910 TGGGTGTAACCTGGTGTCTGAGTCTCTGAGATCGCCACCATGTCAGTGTGAGATGGG 2969  
QY 399 lGluCysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysLysMetGl 919  
Db 2970 GCCCTGACGCGCTCCAGGGAATACATCATCGCTGAGCAGATGCTGTGAGATGGG 3029  
QY 919 yGluAlaLysProSerGlnHisAlaGlyPheValLutIleCysValAlaValCysArgPr 939  
Db 3030 CCATGCCCTCTGCGGAACACCTCCGCGGCGAGTACGCTGTGTATTGGCGAGTGTAAAGCC 3089  
QY 939 oGluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLe 959  
Db 3090 AGAGTTTCATGACGAAGTCCCATCAGCAGTACACTTCTGTAACCTCTCTGTGCTGCTAC 3149  
QY 959 uLysProSerArgGlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLe 979  
Db 3150 CAACCAATCCGAGTCCCGAGTCAAGGAGGAGTATGTTGACCATTTACCGGCAATACCT 3209  
QY 979 uAsnAlaGlySerAsnValValMetPheGlyLysGlnProCysLeuPheHisArgAr 999  
Db 3210 TGGGCTGGGAGCAGCGTGGCAGTCTACCTGGGCAACAGAGCTGGCAGTCTTACGGGAG 3269  
QY 999 gSerProSerTyrIleValCys---AsnThrThrSerSerAspGluValLeuGluMetLy 1018  
Db 3270 GTCAATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3329  
QY 1018 sValSerValGlnValAspArgAlaLysIleHisGlnAspLeuValPheGlnTyrValGl 1038  
Db 3330 TGTTCCTGAGTGTGCGACCGAGCCCATGCTGGATAGCAACCTTGTGAGTACATAGA 3389  
QY 1038 uAspProThrIleValArgIleGluProGluTrpSerIleValSerGlyAsnThrProIl 1058  
Db 3390 TGACCTCGGTCCAGCGCATCGAGCAGTGGAGCATTTGCCAGTGGCCACACACCCCT 3449  
QY 1058 eAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGl 1078  
Db 3450 GACCATCACAGCTTCAACCTGGATGTCTATTCAGGAGCGCAAGGATCCGAGTCAAAATCAA 3509  
QY 1078 yGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAl 1098  
Db 3510 TGGCAAAAGATCTGTCAATGTGTAAAGTGTGTAAAGTGTGTAAAGTGTGTAAAGTGTGTGG 3569

QY	1098	aProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPh	1118	Db	4647	CTTCATGCTATACTGTCATCAAGCAGCAGATGGAGAAGGGCCCATTTGATCCCATCAC	4706
Db	3570	ACCCTCTCTGACCACGACACTACCGCCCTGGCCTGGACACTGTGAACGCCAGATGAGTT	3629	QY	1478	rGlyGluAlaArgTyrSerLeuSerCluAspLysLeuIleArgGlnIleAspTyrLy	1493
QY	1118	eGlyPheIleLeuAsnValGlnSerLeuLeuIleLeuAsnLysThrAsnPhThrTy	1138	Db	4707	GGCCGAGCCCGCTACTCCCTGAGCGAGGACAAGCTCATCCCGCAGCATGATCAGTACAA	4766
Db	3630	TGGATTGTCTTTAAACAATGTCGAATCTTGCTAATTACAACGACACCAAGTTATCTA	3689	QY	1498	sThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLy	1518
QY	1138	rTyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeuLysProGl	1158	Db	4767	GACCTGTATCCTGAAGCTGCTCAACCTGACACGAGAGAACAGTCCAGAGATCCAGTGAA	4826
Db	3690	CTACCCCAACCCGACCTTGAACCTGCTAGCCCTACTGGAGTCTTGGATCAAAAGCCAGG	3749	QY	1518	sIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLy	1538
QY	1158	yThrProIleIleLeuLysGlyLysAsnLeuIleProProValAlaGlyGlyAsnValLy	1178	Db	4827	GGTGTAAACTGTGACACCATCACAGGTCGAAGGACAAGATTCTTGATGCCGTGTATAA	4886
Db	3750	ATCCGCCATCACTCTGAAGGCAAAACCTCTGCCCTCTGAGGG--GCCAA	3806	QY	1538	sAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTrpArgGlnGl	1558
QY	1178	sLeuAsnTyrThrValLeuValGlyGlyProCysThrValThrValSerAspValGl	1198	Db	4887	GAATGTGCCCTATTCCAGCGCGCGAGGAGTGGACATGAGTGTGAGTGGGCCAAGG	4946
Db	3807	ACTCAACTACACTGTGCTCATCGGAGACACCCCTGTGCTGTCAACCGTATCTGAGACCCA	3866	QY	1558	ySerGlyAlaArgMetIleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTr	1578
QY	1198	nLeuLeuCysGluSerProAsnLeuIleGlyArgHisLysValMetAlaArgValGlyGl	1218	Db	4947	CCGGATCGCCCGGCTGCTGCAAGATGAGGACATCACCAACAAGATTGAGGGTGACTG	5006
Db	3867	GCTTCTCTGCCAGCCTCCCAACCTCACCGGCGAGCAAGGTCCTTGGTTCACGTGGCGG	3926	QY	1578	physArgLeuAsnThrLeuAlaHisTyrGlnValProAspGlySerValValAlaLeuVa	1598
QY	1218	yMetGluTyrSerProGlyMetValTyrIleAlaProAspSerProLeuSerLeuProAl	1238	Db	5007	GAAGCGGCTCAACACACTGTATCATATCAGGTGTACAGAGTTCGGTGGTGGCTCTGGT	5066
Db	3927	GATGGTGTCTCGCTGCTCGGTGAGTGTATCTCATCTCAGACAGCTTGTGACCCCTGCCAGC	3986	QY	1598	lSerLysGlnValThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSerAl	1618
QY	1238	alleValSerIleAlaValAlaGlyLeuLeuIleIlePheIleValAlaValLeuI	1258	Db	5067	CCCCAAACAGACCTCCTCTTACAACATCCCTGCTCTGCCAGCATCTCCCGAGCTCCAT	5126
Db	3987	CATCTCAGCATCGCGCGCGCGGAGCCTCTCTCTCATCATCATCATCTGCTCTCAT	4046	QY	1618	aSerLysTyrGluAsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgTh	1638
QY	1258	eAlaTyrLysArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMe	1278	Db	5127	CAGCAGATATGACTCCTCTTACAGGTATACGGGCGCGCCGACAGCCTGCCGTCCCGGGC	5186
Db	4047	TGCCTACAGCCCAAGTCTCGAGAAATAGACCTCACTCTCAAGCGGCTCCAAATCGAGAT	4106	QY	1638	rProMetIleThrProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHi	1658
QY	1278	tAspAsnLeuGluSerArgValAlaLeuGluCysLysGluAlaPheAlaGluLeuGlnTh	1298	Db	5187	CCGATGATACCCCGACCTGGAAAGTGGGTCAAGGTGGCATCTGGTGAAGAACA	5246
Db	4107	GGACAATCTGGAGTCCCGCTGTGGCTTGGAGTGAAGAAAGCTTTTGTGAGTCCAGAC	4166	QY	1658	sGluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLe	1678
QY	1298	rAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspTyrAr	1318	Db	5247	TGACACCGGTGACCAGAGGAGGTGACCCGGGCGAGCAAGATGGTGTCCGAGATCTACCT	5306
Db	4167	GGATATCAATAGTTGACAGTGACTGGACCGCTCAGAAATCCCTTACTTGACACTCG	4226	QY	1678	uThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspAspLeuPheGluTh	1698
QY	1318	gThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLe	1338	Db	5307	GACCCGGCTACTGGCCACCAAGGCCCTCCACAAAGTTTGTGGACGAGTTGTTTGAGAC	5366
Db	4227	TACCTACGCTATGCGAGTCTCTCCCGGCGATCGAGGACCACCCCGCTCTGCGGAGCT	4286	QY	1698	rIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAs	1718
QY	1338	uGluValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLe	1358	Db	5367	CTTGTTCAGCACCTGTCACCCGGGCGAGCGCTCTCCCTGCGCCATCAAGTACATGTTGA	5426
Db	4287	GGAGGTACAGGAACCGGCGAGCAGCAGCTGGAGAAGGCCCTGAAAGCTCTTTGCCAGCT	4346	QY	1718	pPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTr	1738
QY	1358	uIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPh	1378	Db	5427	TTTCTAGATGAGCAGGACAGCAGCATCATGACAGATGTGCGGCACACCTG	5486
Db	4347	CATCAACAACAGGTGTTCTGCTGACCTTATCCCGCACCTCGAGCTCGAGCGCAGTTT	4406	QY	1738	pLysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheVa	1758
QY	1378	eSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLe	1398	Db	5487	GAAAGCAACTGCCCTCCCTCTGCGCTCTGGGTGAACGTGATTAAAGACCCCGCTCGT	5546
Db	4407	CTCCATGCGCGCGCGGCAACCTGGCTTCGCTCATCATGACCGGCTCGAGCGCGCCT	4466	QY	1758	lPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnThrPh	1778
QY	1398	uGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGl	1418	Db	5547	GTTTGTACATCCACAAGGCGCAGCATCACGAGCGCTGCTCTGTGGTGGCCGACACCTT	5606
Db	4467	GGAATATGCACTGATGTCTCAAGCAGCTGCTCTGACCTCATCGATTAAGAACCTGGA	4526	QY	1778	eMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLe	1798
QY	1418	uSerLysAsnHisProLysLeuLeuLeuArgArgThrGluSerValAlaGluLysMetLe	1438	Db	5607	CATGGACTCTTGTCAAGTCAAGCAGCCGGCTGGCAAGGACTCCCCCTCCACACAGCT	5666
Db	4527	GACAAGAACCCCAAGCTGCTACTCCGGAGGACAGAGTCTGTGCTGAAAGATGCT	4586	QY	1798	uLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSerAspIl	1818
QY	1438	uThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLe	1458	Db	5667	GCTCTATGCCAAGGACATCCCGAGCTACAAGAGCTGGGTGGAGAGATACTACGACACAT	5726
Db	4587	GACCAATTTGTTCCCTCTCTCTGCAACAAGTTCTTAAGAGAGTGGCGAGGGGACCACT	4646	QY	1818	eGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluSerAr	1838
QY	1458	uPheSerLeuPheCysAlaIleLysGlnMetGluLysGlyProIleAspAlaIleTh	1478				

```

Db 5727 CGCCAGCTCCCGACATCAGTGACGACGACATGAATGCTACCTCCCGAGCAGTCCCG 5786
QY 1838 gMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGI 1858
Db 5787 CCGCAGCGGTGAGTCAACATGCTGAGTGCCTCAATGAGATCTACTCTTATGTCAG 5846
QY 1858 yLysTyrSerGluGluIleLeuGlyProLeuAspHisAspAspGlnCysGlyLysGlnLy 1878
Db 5847 CAAGTATAGTAGGAGTCACTCGGGGCCCTAGAGCAGGATGAGCAGCAGCGCGGCAGCG 5906
QY 1878 sLeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerLeuAspSer 1894
Db 5907 GCTGGCTATATAGGTGAGCAGCTCATTAATGCGATCTCCATTGAGAGC 5955

RESULT 2
AB007932 6263 bp mRNA linear PRI 13-AUG-1998
LOCUS Homo sapiens mRNA for KIAA0463 protein, partial cds.
DEFINITION AB007932.1 GI:3413887
ACCESSION KIAA0463 protein.
VERSION 1 (sites)
KEYWORDS plus clone:HG0942.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Seki,N., Ohira,M., Nagase,T., Ishikawa,K., Miyajima,N.,
Nakajima,D., Nomura,N. and Ohara,O.
TITLE Characterization of cDNA clones in size-fractionated cDNA libraries
from human brain
JOURNAL DNA Res. 4 (5), 345-349 (1997)
MEDLINE 98116662
REFERENCE 2 (bases 1 to 6263)
AUTHORS Ohara,O.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1997) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@kazusa.or.jp, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
FEATURES
source
1..6263
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="HG0942"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
1..6263
/gene="KIAA0463"
<1..5894
/gene="KIAA0463"
/codon_start=3
/product="KIAA0463 protein"
/protein_id="BAA32308.1"
/db_xref="GI:3413888"
/translation="GFGRLPDSLRAGRSRRPOOPAAAEVDVDRAGTEGOTDVAELSES
CEQCPKVFQMSRPHSRLLTAAPLSMEORPPRALEVDSSRVLLSVNVLAPPAA
GPGQFTSFSESDWTFNHLTVHQGTGAVYVGAINRVYKLTGNLTIOVAHKTPEEDN
KSCYPLIYVPCPELTLTNVKNLLIDYSENRLACGSLYQGVCKLLRDLDFILV
EPHAKHELYLSSVYKGTMTYGVIRSEGEDGKLFITAVDGQDYFFTLSSRKLPRD
ESAMKDELHSDVFSLSLIKIPSDTALVSHDFIYIYFGASGFFVFLTVQPTPEG
VAINSGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRQLQAAVYAKPGDSL
QAFNITISQDDLYFAISKGGOKYVHPPDDSALCAFPRLAIQLKRLQSCYQEGNL
ELNWLKGDVQCTKAPVPIDNNCGLDINQPLGGSTPEVGLTITTSRDRMTSVASIV
YNGSVYVFPVSESGKLTIRADPGPHGGVQYEMVSVLKDGPSPILRDMAFISIDQRLYV
MSERQYTRVPVESEGYTTTCGSGSDGPHCGWALHNMCSRRDKCQAWEPNRYLXV
ISQCVSLAVPHSSISYSEHSRLSLVSDAPDLSAGIACAFGNLTVEVGQVSGSVFC
ISPGKDVPIPLDDQWGLQLRSKETGKIFVSTEFKYNCSAHOLCLISCVNSAFR
CHWCKYRNLTCHDPTTCSQEGRNINISDPCQLVPTPEIILIPVGEYKPTILKARNLQ
PQSGQGYECVLIQQAIRHVPALRPNSSVQCNSSIQYDGMDSINLAVDFVWNG
NFITIDNPQDLKVLHYKCAAQRESGLCLKADRFECCGWCSEGRRCRTLHQHCTSPSSPW

```

```

LDWSSHNVKCSNPQITEILTVSGPPEGGTRVTIHGVNGLGDFSEIAHHVQVAGVPCTP
LPGEVIIAEQIVCEMHALVGTSTGVPVRLCIGCEKPEFMTKSHQOQTFYVNPVSLNP
IRGPSGGTMTVITGHYGAGSSVAVILGNOTCEFYGRSMSEIVCVSPSSNGLGVPV
VSVSDRAHVDNSNQFEXIDDPVORIEPEWSIASGHTPLTITGFNLDVIOEPRIK
ENGRESVNVKVVNTTITLCLAPSLTIDYRPGLDVPERDEFGFVNNVQSLIYNDT
KFIYTPNTELLSPTGLVDQKPSPIILKGRNLCPASGAGKARLNTVILIGETPCAVT
VSEFOLCEPPNLTGQHKVMHVHVGWVSPGSVSDLSLLTLPALVISAAGSLLLI
IVIVLIAYKRKSRENDLTLRLOQMNDLSEKVALECKEAEALQDTNELTSLDLR
SGIPLDYRTYAMRVLEIDHPVLEVEGVQGHVEKALKLFAQLINNKVFLLT
FIRTELOSFSMRDRGNVASIIMTGLQGRLEVATDVLKQLLSLDLXKLNKHPKL
LLRRTESVAEKMLTNWPAFLHKLKECAGEPLMLYCAIKOOMEKCPIDAITGEARY
SLSDEKLIRQOIEYKTLINCVNPDNENSPFVPLVNCNTITQVKEKLDVAYKNVP
YSRPRAVMDLENRQRIARVLOQEDITTKIEGDKRNLMTLMHVQVSDRSVVALVP
KQTSYNIPASASISRTSISRYSDSFRYTGSPDSLRAPMITPDLSEGVKVMHLVKN
HDHGDKGEGDRGSKSEIYLRLLATKTLOKFVDDLPETLFTVHRGSALPLAIKY
MFDLEDOADRHSIHDTDVRHWTKNCLPLRFWVANIKNPOFVEDIHKGSITDACLVS
VAFQFMDSCSTSEHRLGKXPSNKLKYAKDIPYSKSWERYADIADKLPAISODNNA
YLAQSRHLHAFEFNMLSALNEIYSVSKYSEELIGALEODEQARRORLAKYKVEQLINA
MSIES"
BASE COUNT 1359 a 1890 c 1739 g 1275 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 6263
Score: -6790.00 Matches: 1280
Percent Similarity: 80.18% Conservative: 241
Best Local Similarity: 67.47% Mismatches: 358
Query Match: 67.97% Indels: 20
DB: 9 Gaps: 8
US-09-964-956-13 (1-1896) x AB007932 (1-6263)
QY 5 ProTepAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMetGlySerSerThr 24
Db 238 CCCTCGAGGTGACAGCCGCTCTGTG-GTCCGTGCTCTCAGT-GGTCTGGTGTGCTGCT--- 292
QY 25 LeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSer-PheValThrPheAr 44
Db 293 -----GGCCCCCAGCAGCGCGCATGCTCAGTTCAGCAGCCTTCCA 334
QY 44 gGlyGluProAlaGlu---GlyPheAsnHisLeuValValAspGluArgThrGlyHisII 63
Db 335 CTCTGAGATCTGTGACTGGACCTTCAACACACTTGACCGTCCACAGGAGCGGGGCGGT 394
QY 63 eTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValTh 83
Db 395 CTATGTGGGGCCATCAACCGGGTCTATAAGCTGACAGCAACCTCACCATCCAGGTGGC 454
QY 83 rHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIleValGlnTh 103
Db 455 TCATAAGACAGGGCCAGAGAGAGAGACAAAGTCTTGTACCCGCCCTCATCTGTCAGCC 514
QY 103 rCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIleAspTyrLy 123
Db 515 CTGACGGAAGTGTCTACCCCTCACCACCAATGTCAACAAGCTGCTCATATTGACTACTC 574
QY 123 sGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeuLeuArgLe 143
Db 575 TGAGAACCGCTGTGCTGGCGCTCTACAGGCGCTCTACAGGGGGTCTGCAAGCTGCTCGGCT 634
QY 143 uGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeuSerGlyVa 163
Db 635 GGATCACCTCTTCATCCTGGTGGAGCCATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 694
QY 163 lasnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAspLysLe 183
Db 695 CAACAAGACGGGACCATGTATCGGGGTGATTGTGGCGCTCTGAGGGTGGAGGATGGCAAGCT 754
QY 183 uPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSerArgLy 203
Db 755 CTTTCATCGCAGCGCTGTGGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 814
QY 203 sLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheVa 223

```

Db 815 CTTGCCCCGAGACCTGAGTCTCAGCATGCTCGACTATGAGTACACACGATTTGT 874  
Qy 223 lAlaSerMetIleLysIleProSerAspThrPheThrIleProAspPheAspIleTy 243  
Db 875 CTCCTCTCTCATCAAGATCCCTTCAGACACCTGGCCCTGCTCTCCCACTTTGACATCTT 934  
Qy 243 rTyrValTyrGlyPheSerSerGlyAsnPhValTyrPheLeuThrLeuGlnProGluMe 263  
Db 935 CTACATCTACGGCTTGTCTAGTGGGCTTGTCTACTTTCTCACTGTCCAGCCGAG-- 992  
Qy 263 tValSerProGly-----SerThrThyLysGluGlnValTyrThrSerLysLe 280  
Db 993 ----ACCCCTGAGGTTGGCCATCACTCCGCTGGAGACCTCTCTACACCTCAGCAT 1048  
Qy 280 uValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCy 300  
Db 1049 CGTGGGCTCTGCAAGGATACCCCAAGTTCACCTACATACGCTGCTCCCTGGCCCTG 1108  
Qy 300 sGluArgSerGlyValGluTyrArgLeuGluAlaAlaTyrLeuSerLysAlaGlyAl 320  
Db 1109 CACCGGCGGGGTGGAAATACCCCTCTCTCAGCTCTTACCTGGCCAAAGCTGGGA 1168  
Qy 320 aValLeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSe 340  
Db 1169 CTCAGTGGCCAGGCCCTCAATATACACGAGCAGGACGATGCTCTTTGCCATCTTCTC 1228  
Qy 340 rLysGlyGlnLysArgLysMetLysSerLeuAspLysSerAlaLeuCysIlePheIleLe 360  
Db 1229 CAAAGGCGAAGAGATATACACCCCGCCGATGACTGCGCCCTGTGTGCCCTCCCTAT 1288  
Qy 360 uLysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyTh 380  
Db 1289 CCGGCCATCACTGACAGATCAAGGGGCGCTACAGTCTCTACACGAGGCGAGGCAA 1348  
Qy 380 rLeuAspLeuAlaThrLeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIl 400  
Db 1349 CCTGAGCTCAACTGGCTGTGGGGAAGGAGCTCCAGTGCACCAAGGCGCTGCCCCAT 1408  
Qy 400 eAspAspAsnPhcCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValAr 420  
Db 1409 CGATGATAACTCTGTGACGTGGACATCAACCAACCCCTGGAGGCTCAACTCCAGTGA 1468  
Qy 420 gGlyIleProValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrValTy 440  
Db 1469 GGGCTGACCTGACACACAGCAGGAGCGCATGACCTCTGTGGCTCTCTAGCTTTA 1528  
Qy 440 rLysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLysLysIleArgVa 460  
Db 1529 CAACGGCTACAGCGTGGTTTTTGTGGGACTAAGAGTGGCAAGCTGAANAAGATTTCGGC 1588  
Qy 460 lAspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGlnValVal--AspProGl 479  
Db 1589 CGACGGTCCCCCATGTTGGGTCAGTACGAGATGGTCTCTGTCTGCTCAAGCAGGAAG 1648  
Qy 479 yProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGl 499  
Db 1649 CCCCATCTCGGGACATGGCCCTTCCATTGATCAGGCTACCTACCTGCTAGCTGCTGA 1708  
Qy 499 uArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrGlnSerCysGlyGlyCy 519  
Db 1709 GAGACAGGTACACAGGTCCTGGGTGATGTCAGTACGTATACGACTTGTGGGAGTG 1768  
Qy 519 sLeuGlySerGlyAspProHisCysGlyTrpCysValLeuLeuHisAsnThrCysThrArgLy 539  
Db 1769 CTTGAGCTCTGGGACCTCACTGTGGTGGTGTGCCCTGCTGCAACATGCTGCTCCGCGAG 1828  
Qy 539 sGluArgCysGluArgSerLysGluProArgPheAlaSerGluMetLysClnCysVa 559  
Db 1829 GGACAAATGCCAACAGGCTGGGAACCTAATCGATTTGCTGCCAGCATCAGCCAGTGTGT 1888  
Qy 559 lArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuValLe 579  
Db 579 lArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuValLe 579

Db 1889 GAGCCTTGCAGTGCATCCAGCAGCAGCTCTCAGTATCTGAGCAGCAGCGGTTGCTAGCCT 1948  
Qy 579 ucluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSe 599  
Db 1949 GGTAGTGTAGTGTCTGATCTATCTGGGTATCGCTGTGCTTTGGGAACCTGAC 2008  
Qy 599 rGluMetAspGlyLeuValValGlyAsnGlnIleGlnCysTyrSerProAlaAlaLysGl 619  
Db 2009 AGAGTGGAGGCGAGGTCTCCGGAGCCAGGTCATCTGCATCTACCTGGGCCCAAGGA 2068  
Qy 619 uValProArgIleLeuThrGluAsnGlyAspHisValValGlnLeuLeuLysSe 639  
Db 2069 TGTCCCT---GTCATCCCGCTGATCAAGACTGTTGGCTGGAGCTAGAGTGGTCT 2125  
Qy 639 rLysGluThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerValHi 659  
Db 2126 CAAGGAGACGGAAGATATTTGTCACACCGAGTTCAAGTTTTTACAACTGCGATGCCCA 2185  
Qy 659 sAsnSerCysLeuSerCysValGluSerProTyrArgCysHisTrpCysLysTyrArgHi 679  
Db 2186 CCAACTGTGCTGCTGCTGTCACACGCGCTTCCGCTGCAATTTGGTGAAGTACCGCAA 2245  
Qy 679 sValCysThrHisAspProLysThrCysSerPheGlnGluGlyArgValLysLeuProGl 699  
Db 2246 CCTGCTCATCATGACCCCAACCTGCTCTTCCAGGAGGCGGATCAATATTCAGA 2305  
Qy 699 uAspCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysPr 719  
Db 2306 GGACTGTCCCGCTGCTGCTGCCACAGAGGAGATCTTGATTCAGTGGGAGGTAAAGCC 2365  
Qy 719 oIleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCy 739  
Db 2366 AATCACCTTAAGGCGCGAATCTGCCACGCGCAGTCCGCGCAGCGAGGCTATGATG 2425  
Qy 739 sIleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArgPheAsnSerSerSe 759  
Db 2426 TGTCCTCAACATACAGGAGCCATCCACCGGTCCCGCTCTGGCTTCAACAGCTCCAG 2485  
Qy 759 rValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsnLeuProVa 779  
Db 2486 CGTTCACTGTCAACAGCTCGTACCATGATGATGGCATGGCATGACATCGCAATCTGGC 2545  
Qy 779 lGluLeuThrValValTyrAsnGlyHisPheAsnIleAspAsnProAlaGlnAsnLysVa 799  
Db 2546 GGATTTTCGTGCTGTGGAAACGCAATTCATCATTTGACAACCTCAGGACCTCAAGT 2605  
Qy 799 lHisLeuTyrLysCysGlyAlaMetArgGluSerCysGlyLeuCysLeuLysAlaAspPr 819  
Db 2606 CCATCTCTCAAGTTCACGCCAGCGGAGAGTGGCGCTCTGCTCAAGGCGGAGCCG 2665  
Qy 819 oAspPheAlaCysGlyTrpCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysPr 839  
Db 2666 GAAGTTTGTGCTGTGGTGTGCGCGCGAGCGAGTGCACCTCCACCGACCTGATC 2725  
Qy 839 oAlaGlnGluSerGlnTrpLeuGluLeuSerGlyAlaLysSerLysCysThrAsnProAr 859  
Db 2726 CAGCCCTTCAGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2785  
Qy 859 gIleThrGluIleIleProValThrGlyProArgGluGlyGlyThrLysValThrIleAr 879  
Db 2786 AATCACCAGATTTTGACGGTGTGTCACCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 2845  
Qy 879 gGlyGluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyVa 899  
Db 2846 TGGCGTGAACCTGGCTGGAGTTCTCCGAGATCCGCCACCATGCTGAGGTTGCTGCTGG 2905  
Qy 899 lGluCysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMetGl 919  
Db 2906 GCCCTGCACGCCCTCCCGGGGAAATACATCATCGCTGAGCAGATTTGCTGTGAGATGG 2965  
Qy 919 yGluAlaLysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgPr 939  
Db 2966 CCATGCGCTCTGGGAACACCACTCCGGGCCAGTACGCTGTGTATTTGGCGAGTGTAA 3025

1298	Qy	rAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspTyrAr	1318
4103	Db	GGATATCAATCAGTGTGACCACTGACCTGGACCGCTCAGGAATCCCTTACCTTGGACTATCG	4162
1318	Qy	gThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLe	1338
4163	Db	TACCTACGCTATGCGAGTCTTGTTCCCGGGCATCGAGGACCACCCGCTCTCGCGGAGCT	4222
1338	Qy	uGluValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLe	1358
4223	Db	GGAGGTACAAAGAAACCGGGCAGCAGCAGCTGGAGAAAGCCCTGAAGTCTTGTGCCCACTG	4282
1358	Qy	uIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPhe	1378
4283	Db	CATCAACAAACGAAGTGTCTCTGACCTTCATCCGACCCCTGGAGTGCAGCCAGTTT	4342
1378	Qy	eSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLe	1398
4343	Db	CTCCATGCGCGACCGGGCAACGTGGCTTCGCTCATGACCGCGCTGCAGGGCCGCT	4402
1398	Qy	uGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGI	1418
4403	Db	GAATATGCCACTGATGTCCCTCAACAGCTGCTCTGACCTCATCATGATAAGAACCTGGA	4462
1418	Qy	uSerLysAsnHisProLysLeuLeuLeuArgArgThrGluSerValAlaGluLysMetLe	1438
4463	Db	GAACAAGAACCACCCCAAGCTGCTACTCCGGCAGGACAGACTCTGTGGCTGAAAGATGCT	4522
1438	Qy	uThrAsnThrPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLe	1458
4523	Db	GACCAATTTGGTTTCGCTTCTCTGCAACAGTTCTCAAGGAGTGCAGCGGGAGCCACT	4582
1458	Qy	uPheSerLeuPheCysAlaIleLysGlnGlnMetGluLysGlyProIleAspAlaIleTh	1478
4583	Db	CTTCATGCTATACTGTGCCATCAACAGCAGATGGAGAGGGCCCCCATTTGATGCCATCAC	4642
1478	Qy	rGlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLy	1498
4643	Db	GGGGAGCGCGCTACTCCTCTGAGCGAGGACAAGCTATCCGGCAGCAGATCGAGTACAA	4702
1498	Qy	sThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLy	1518
4703	Db	GACCTGTACTGAACTGCGTCAACCTTGACACGACGACAGTCCAGAGATCCCAGTGAA	4762
1518	Qy	sIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLy	1538
4763	Db	GGTGTAAACTGTGACACCATCACACAGTCAAGGACAAGATTTCTGATGCCGTGTAA	4822
1538	Qy	sAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTrpArgGlnGI	1558
4823	Db	GAATGTGCCCTATTCCACGCGCCGAGGGCAGTGGACATGGACTTGGAGTGGCCCAAGG	4882
1558	Qy	ySerGlyAlaArgMetIleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTr	1578
4883	Db	CCGGATCGCCCGGTGCTGCTGCAAGATGAGGACATCACCACCAAGATTGAGGGTGACTG	4942
1578	Qy	plyArgLeuAsnThrLeuAlaHisTyrGlnValProAspGlySerValValAlaLeuVa	1598
4943	Db	GAACGGCTCAACACACTGATGTCAATTTATCAGTGTCTCAGCAGGTCGGTGGTCTGCT	5002
1598	Qy	lSerLysGlnValThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSerAl	1618
5003	Db	CCCCAACAGACCTCTCTCTACACATCCCTGCTGCGCAGCATCTCCCGGACCTCCAT	5062
1618	Qy	aSerLysTyrGluAsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgTh	1638
5063	Db	CAGCAGATATGACTCTCTTTCAGGTATACGGGCAGCGCCGACAGCTCGCGTCCCGGC	5122
1638	Qy	rProMetIleThrProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHi	1658
5123	Db	CCCGATGATACCCACGACCTGGAAGATGGGGTCAAGTGTGGCATCTGGTGAAGAACC	5182
1658	Qy	sGluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLe	1678

Db	5183	TGACCAGGTGACCAAGAGGAGGTGACCGGGCAGCAAGATGGTGTCCGAGATCTACCT	5242	
Qy	1678	uThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspLeuPheGluTh	1698	
Db	5243	GACCGGCTACTGGCCACCAAGGACCCCTGCAGAAAGTTTGTGGACGACTTGT	5302	
Qy	1698	rIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaLysTyrMetPheAs	1718	
Db	5303	CTTGTTACGACTGTGACCGGGGCGAGCGCTCTCCCTGGCCATCAAGTACATGTTGA	5362	
Qy	1718	pPheLeuAspGluGlnAlaAspLysHisGlyLeuHisAspProHisValArgHisThrTr	1738	
Db	5363	TTTCTAGATGAGCAGGACAGCAGCAGCATCCATCACAGATGTCCGGCACACCTG	5422	
Qy	1738	pLysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheVa	1758	
Db	5423	GAAGAAGCAACTGCCCTCCCTCTGCGCTTCTGGGTGAACGTGATTAAAGAACCCCACTTCGT	5482	
Qy	1758	lPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValAlaGlnThrPh	1778	
Db	5483	GTTTGACATCCACAAAGGCGACATCACGAGCGCTGCTCTGTGTGGGCCACACCTT	5542	
Qy	1778	eMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLe	1798	
Db	5543	CATGGACTCTTGTTCACGTACAGCAGCAGCGGTGGGCAAGGACTCCCTCTCCAAAGCT	5602	
Qy	1798	uLeuTyraLysAspIleProSerTyraLysAsnTrpValGluArgTyrTyrSerAspIl	1818	
Db	5603	GCTCTATGCCAAGAGCATCCCGAGTACAGAGCTGGGTGGAGATACTACGACAGCAT	5662	
Qy	1818	eGlyLysMetProAlaLysSerAspGlnAspMetAsnAlaTyraLeuAlaGluGlnSerAr	1838	
Db	5663	CGCCAAGCTCCAGGCATCAGTGACGAGCATGAATGCTACCTCGCGGAGCAGTCCCG	5722	
Qy	1838	gMethHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGl	1858	
Db	5723	CCTGCACCCGTTGGAGTTCAACATGCTGAGTGCCCTCAATGAGATCTACTCTATGTACG	5782	
Qy	1858	yLysTyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCysGlyLysGlnLys	1878	
Db	5783	CAAGTATAGTGGAGGACTCATCGGGCCCTAGAGCAGGATGAGCAGGCGCGGCGAGCG	5842	
Qy	1878	sLeuAlaTyraLysLeuGluGlnValIleThrLeuMetSerLeuAspSer	1894	
Db	5843	GCTGGCTTATAAGGTGGAGCAGCTCATTAATGCCATGTCCATTGAGAGC	5891	
RESULT 3				
AX305728				
LOCUS	AX305728	6730 bp	DNA	linear
DEFINITION	Sequence 479 from Patent WO0188188.			
ACCESSION	AX305728			
VERSION	AX305728.1	GI:17645146		
KEYWORDS	house mouse.			
SOURCE	Mus musculus			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1			
TITLE	Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.			
JOURNAL	Method for examining ischemic conditions			
FEATURES	Patent: WO 0188188-A 479 22-NOV-2001;			
Source	School Juridical Person Nihon University (JP)			
Location/Qualifiers	1..6730			
BASE COUNT	1656 a 1650 c 1740 g 1482 t			
ORIGIN	/organism="Mus musculus"			
	/db_xref="taxon:10090"			
Alignment Scores:				
Pred. No.:	0	Length:	6730	
Score:	6765.00	Matches:	1268	

Percent Similarity:	80.24%	Conservative:	239
Best Local Similarity:	67.52%	Mismatches:	357
Query Match:	67.72%	Indels:	14
DB:	6	Gaps:	8
US-09-964-956-13 (1-1896) x AX305728 (1-6730)			
Qy	25	LeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPheArg	44
Db	628	CTGTAGTCGTGGGACACACAGGTATGCGCTCAG-----TACAGCAGCTTTCCAC	675
Qy	45	GlyGluProAlaGlu---GlyPheAsnHisLeuValValAspGluArgThrGlyHisIle	63
Db	676	TCTGAGAAATCGTACGTGAGCTTTTCAACCACTTCTACCTGCTACACGAGAACAGGGCGTGTG	735
Qy	64	TyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThr	83
Db	736	TATGTGGGGGCTATCATCGTGTCTACAGTTGACTGGCAACCTCACCTACCTACCTGGCT	795
Qy	84	HisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIleValGlnThr	103
Db	796	CACAGACAGGGCCAGAGAGGACACAGGCTTGCTACCCACCCCTCATTTGTACAGCCC	855
Qy	104	CysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuLeuAspTyrLys	123
Db	856	TGCAGTGAAGTGTACACTCACCAACAATGTCAACAACACTACTGATCATTTGACTACTCT	915
Qy	124	GluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeuLeuArgLeu	143
Db	916	GAGAATCGCTGCTGGCTGGGAGCCCTCTACCGAGGTGTTTGCAGGCTCTCTCGACTA	975
Qy	144	GluAspLeuPheLysLeuGlyGluProTyraHisLysLysGluHisLysLeuSerGlyVal	163
Db	976	GATGACCTCTTCTATCTGTTGGAGCCATCCACAGAAGAACACTACTTGTCTCCAGTGT	1035
Qy	164	AsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAspLysLeu	183
Db	1036	ATAAAGACAGGACCACTGATGTGTGATTGTGGCTCTGAGGGGAGAGATGGCAGCTT	1095
Qy	184	PheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSerArgLys	203
Db	1096	TTTATCGGCACCTCTGTGGATGTCACAGCAGGATTTCTCTACTCTGTCTCCAGCGCAAG	1155
Qy	204	LeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyraValPheHisAspGluPheVal	223
Db	1156	CTCCCGCTGACCTTTCAGCAATGCTTGACACTCTGAGCTTCCAGCTTCCAGCTGTC	1215
Qy	224	AlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAspIleTyr	243
Db	1216	TCCCTCCCTCATCAGATTCCTCTGACACCTAGCCCTGGTCTCTCACTTCGACATCTTC	1275
Qy	244	TyrValTyraGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMet	263
Db	1276	TACATCTATGCTTTCGACGTGGGGGCTTGTCTACTTTCTCTCACTGCTCCAGCCAGAG	1332
Qy	264	ValSerProGly-----SerThrThrLysGluGlnValTyrThrSerLysLeu	280
Db	1333	---ACCCCTGACCGCTGGCCATCAATCTACGCTGAGACCTCTTCTATACCTCAAGAATT	1389
Qy	281	ValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCys	300
Db	1390	GTGCGTCTCTGCAAGGATGACCCCAAGTTCACCTCTATGTCTCCCTTTTGGCTGC	1449
Qy	301	GluArgSerGlyValGluTyraArgLeuLeuGlnAlaAlaTyraLeuSerLysAlaGlyAla	320
Db	1450	ACAGTGTCTGGGTGGAATATCGCTTCTGACGAGCAGCTTACCTTGCAAAACCCAGGAA	1509
Qy	321	ValLeuGlyArgThrLeuGlyValHisProAspAspLeuPheThrValPheSer	340
Db	1510	GCTCTAGCTCAGCGCTTCAACATCAGCAGGACGAGATGCTCTGTTGGCCTCTTTTCC	1569
Qy	341	LysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeu	360

1570	Db	AAAGGGCAGAAAGCAGTACACACACCCCTGTAGTACTCTCCCTCTGTGTCTCTCCCATC	1629
361	Qy	LysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThr	380
1630	Db	CGGCAATCAACTTGCAAAATCAAGGAGCGGTTCAGTCTCTACCAAGGAGGGAAC	1689
381	Qy	LeuAspIleAlaTrpIleuLysValLysAspIleProCysSerSerAlaLeuThrIle	400
1690	Db	TTGGAGCTCAACTGGCTGTGGGAAAGGATGTGCAGTGCACCAAGGCGCTGTCCCAATC	1749
401	Qy	AspAspAsnPhcCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArg	420
1750	Db	GATGATAACTTCGCGCCCTGGACATACACCAAGCTCTGGGAGGCTCCACTCTGTGGAG	1809
421	Qy	GlyIleProValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrValTyr	440
1810	Db	GGATGACCCCTGTATACCAACACGAGGACCGCTGACCTCTGTGGCCCTCTATGTTTAC	1869
441	Qy	LysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLysIleAerGVal	460
1870	Db	AATGGGTACAGTGTGGTTTTGTGGGACTTAAGAGTGGCAAGCTGAAGAAGATTCGAGCT	1929
461	Qy	AspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGlnValVal---AspProGly	479
1930	Db	GATGGTCCCCCATGGTGGGGTCCAGTATGAGATGGTCTCTGTGTCAAGATGGGAGC	1989
480	Qy	ProValLeuArgAspMetAlaPheSerLysAspHisGlnGlnLeuTyrIleMetSerGlu	499
1990	Db	CCAATCTCGGACATGGCTTCTCCATCAATCAGCTATACCTATATGTCATGTCGTGAG	2049
500	Qy	ArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrGlnSerCysGlyGluCys	519
2050	Db	AGACAGTCAACCAAGGTCCTGTGTAATCATGTGAACAGTATACAACCTTGTGGAGAGTGT	2109
520	Qy	LeuGlySerGlyAspProHisCysGlyTrpCysValLeuHisAsnThrCysThrArgLys	539
2110	Db	CTAAGCTCAGGGATCTCACTGTGGTGTGTGGCCCTGCACAACTGTCCTCCCAAGA	2169
540	Qy	GluArgCysGluArgSerLysGluProArgArgPheAlaSerGluMetLysGlnCysVal	559
2170	Db	GACAAATGCCAACGGCCCTGGGAAGCAAAATCGATTGCTGCCAGTATCACCCAGTGCATG	2229
560	Qy	ArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuLeuValLeu	579
2230	Db	AGCCTTGGAGTACACCCCAACAGCATCTCTGTGTGCAGATCACAGCGCGTCTCAGCCGTG	2289
580	Qy	GluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSer	599
2290	Db	GTTGTGAATAGTGTCCCAACCTCTCTGAAGGTATTGCTGTGCCCTTGGGAATCTGACT	2349
600	Qy	GluMetAspGlyLeuValValGlyAsnGlnIleGlnCysTyrSerProAlaAlaLysGlu	619
2350	Db	GAGGTGGAGGCAGGTATCTGGGAGTCAAGTCAATCTGCATCTCACCTGGACCCCAAGGAT	2409
620	Qy	ValProArgIleIleThrGluAsnGlyAspHisHisValValGlnLeuGlnLeuLysSer	639
2410	Db	GTCCTCT---GTCATCCTCTGGATCAAGACTGTTTGGCCCTAGAGCTGCAGCTGAGATCC	2466
640	Qy	LysGluThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerValHis	659
2467	Db	AAAGAGACAGGAAGAATCTTTGTGCACACCGGAATTCAAAGTTCATACTGCATGCCAC	2526
660	Qy	AsnSerCysLeuSerCysValGluSerProTyrArgCysHisThrCysLysTyrArgHis	679
2527	Db	CAACTGTGCTGTCCCTGTGTATACAGCGCTCTCCGTGCCATTTGGTGAAGTACCGTAAC	2586
680	Qy	ValCysThrHisAspProLysThrCysSerPheGlnGluGlyArgValLysLeuProGlu	699
2587	Db	CTCTGCACACATGACCCCACTACCTGTCTCTCCAGGAAGCAGGATCAATGTTTCCAGAG	2646
700	Qy	AspCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysPro	719
2647	Db	GACTGTCTCCCAAGTCTGTGCCACGGAGGAGATTTCTGATCCAGCTTGGGAAGTAAACA	2706

Qy	720	IleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCys	739
Db	2707	ATCACCCCTTAAGGCCCGAAACCTGCCCCAGCCAGCTGCGCCAGCGAGGCTACGAGTGT	2766
Qy	740	IleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArgPheAsnSerSerSer	759
Db	2767	GTGCTCAGCATCAAGGGGTCCACCGGGTCCCTGCCTCGCTTTCAACAGTTCCAGT	2826
Qy	760	ValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsnLeuProVal	779
Db	2827	GTGCAGTGCACAAACAGCTCGTACCATATGATGGGATGGACATCAGCAACCTAGCAGTG	2886
Qy	780	GluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsnProAlaGlnAsnLysVal	799
Db	2887	GACTTTGCTGTAGTATGGAATGGCAACTTCATTATTGACAACCTCAGGACCTCAAGATA	2946
Qy	800	HisLeuTyrLysCysGlyAlaMetArgLusSerCysGlyLeuCysLeuLysAlaAspPro	819
Db	2947	CATCTCTACAAGTGTGCGCTCCAGCGGGAAGGTGTGTCTCTGCCTCAAGGTGACCA	3006
Qy	820	AspPheAlaCysGlyTyrCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysPro	839
Db	3007	AAGTTTCAGGTGTGCTGGTGGTGGTGGCGCAGATGTACCTTCCACAGCACTGCCCC	3066
Qy	840	AlaGlnGluSerGlnTrpLeuGluLeuSerGlyAlaLysSerLysCysThrAsnProArg	859
Db	3067	AGCACTTCTAGCCCTGGCTTGACTGTCAGCCACAATGTCAAGTGTTCACACCCCAA	3126
Qy	860	IleThrGluIleLeuProValThrGlyProArgGluGlyGlyThrLysValThrIleArg	879
Db	3127	ATCACAGAGATTTTGACAGTATCAGGACCACCTGAAGGAGGAGCTCGTGTGACCATCCAT	3186
Qy	380	GlyGluAsnLeuGlyLeuGluPheArgAspIleAsSerHisValLysValAlaGlyVal	899
Db	3187	GGCTGGAACCTGGCTGGACTTCTCGAGATTGCTCACCATGGCAGGTGGTGGAGTG	3246
Qy	900	GluCysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMetGly	919
Db	3247	CCCTGACACCTATCCCAGGGGAATACATCATCGCTGAGCAGATCGTGTGAGATGGC	3306
Qy	920	GluAlaLysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgPro	939
Db	3307	CATCCGCTTATAGGTACACATCTGGCCCTGTGCGCTGTGCATTTGGGGAATGCAAGCCA	3366
Qy	940	GluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeu	959
Db	3367	GAGTTTCATGACCAGTCCCACACGACGATACTTTTGTGAATCCTCTGTGCTGTCACTC	3426
Qy	960	LysProSerArgGlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeu	979
Db	3427	AGCCCGATCCGGGACACGAGTACAGAGGTACCATGTTGACCATCAGAGCCATTACCTT	3486
Qy	980	AsnAlaGlySerAsnValValMetPheGlyLysGlnProCysLeuPheHisArgArg	999
Db	3487	GGTCTGGGACAGTGTGGCAGTGTACCTGGGCAATCAGACCTGTGAAATTCATTGGGAGG	3546
Qy	1000	SerProSerTyrIleValCys---AsnThrThrSerSerAspGluValLeuGluMetLys	1018
Db	3547	TCAATGAATGAGATTTGATGTTTTCACCCCATCATCCCAATGGACATGAGCAGCCCT	3606
Qy	1019	ValSerValGlnValAspArgAlaLysIleHisGlnAspLeuValPheGlnTyrValGlu	1038
Db	3607	GTCTCCCTGAGTGTGCACAGACCCGGGTGGATAGCAGTCTGCAGTTCGAGTATATAGAT	3666
Qy	1039	AspProThrIleValArgIleGluProGluTrpSerIleValSerGlyAsnThrProIle	1058
Db	3667	GACCCACGGGTCCAACATTTAGGCCAGAGTGGAGTATCATAGTGGGACACACACCCCTA	3726
Qy	1059	AlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGly	1078
Db	3727	ACCATCAGCGTTCACACTTGGATGTCTATTCAGGAGGCCAGGGTCCGAGTCAAAATTAAT	3786

QY	1079	GlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAla	1098	Db	4864	ACTAAGTGGTTGGTCTTCTTCTACAAAGTCTCTGAAGAGTGTCTGGGAACCACTC	4923
Db	3787	GGCAAAAGTCTGCAATGATGACAGTGTGAACACAAACACCACTGCTGGCA	3846	QY	1459	PheSerLeuPheCysAlaIleLysGlnMetGluLysGlyProIleAspAlaIleThr	1478
QY	1099	ProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPhe	1118	Db	4924	TTTATGCTATATCTGTCATCAATCAAGCAGCAGATGGAAGAGCCCATTTAGCTATTA	4983
Db	3847	CCCTCTCTGACCACTGACTACCGCCAGGTCTGACACTGTGGACGCCAGATGAGTT	3906	QY	1479	GlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLys	1498
QY	1119	GlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsnPheThrTyr	1138	Db	4984	GTTGAGGCCCGATCTCCCTGAGTGAAGCAAGCTCATCCGCGCAGAGATCGAGTATA	5043
Db	3907	GGATTCTCTTAAACAATGTTCAATCTTACTATCTATAACGACCAACCAAGTTCTAT	3966	QY	1499	ThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLys	1518
QY	1139	TyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeuLysProGly	1158	Db	5044	ACTCTGATCCTCAACTGTGTCAACCTGTCAATGAGAACAACGCCAGAGATCCCAAGT	5103
Db	3967	TACCCCAACCAACGTTTGAACCTGCTCAGCCCACTGGAATCTTGATCAGAACCGC	4026	QY	1519	IleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLys	1538
QY	1159	ThrProIleLeuLysGlyLysAsnLeuIleProProValAlaGlyGlyAsnValLys	1178	Db	5104	GTACTAACTGTGACACCATCACTCAAGTCAAGAGAGAAGATCTCGATCGGTATA	5163
Db	4027	TCACCCATCACTCTGAAGGCAAAATCTCTGCTCTGCTGCTGGAGG---GCCAAA	4083	QY	1539	AsnValProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTyrArgGlnGly	1558
QY	1179	LeuAsnTyrThrValLeuValGlyGlyLysProCysThrValThrValSerAspValGln	1198	Db	5164	AATGTCCCTACTCCAGCGCCCAAGGGCTGTGGACATGGATCTGGAGTGGCGCAAGC	5223
Db	4084	CTCAACTACACAGTAATGATTGGAGACACCTTGTACAGTCACCTGTGTGAGACAG	4143	QY	1559	SerGlyAlaArgMetIleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTrp	1578
QY	1199	LeuLeuCysGluSerProAsnLeuIleGlyArgHisLysValMetAlaArgValGlyGly	1218	Db	5224	CGGATTGCCAGAGTGGTGTTCAGGACGACAGACATTACCACCAAAATAGAGGTG	5283
Db	4144	CTGCTTTTGAACCTCCCACTCACAGGCGACAGCAAGTTCATGTTCCAGTGGCGGG	4203	QY	1579	LysArgLeuAsnThrLeuAlaHisTyrGlnValProAspGlySerValValAlaLeuVal	1598
QY	1219	MetGluTyrSerProGlyMetValTyrIleAlaProAspSerProLeuSerLeuProAla	1238	Db	5284	AAGCGGCTTAACACACTGATGCATTACAGGTGTACAGAGATCCGTGGTGGCTGGT	5343
Db	4204	ATGGTGTCTCACTCCCTGCTGAGCGTCATCTCCGACAGCTGTGTGACCTGGCAGC	4263	QY	1599	SerLysGlnValThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSerAla	1618
QY	1239	IleValSerIleAlaValAlaGlyGlyLeuLeuIleIlePheIleValAlaValLeuIle	1258	Db	5344	CCTAAGCAGACCTCTCTCTACAACTCCCTCCCTGCCAGCATCTCTCGACATCCAT	5403
Db	4264	ATCATCAGATCGACGCTGGTGAACCTCTCTCTTATCATCTCATCTGTCATCTG	4323	QY	1619	SerLysTyrGluAsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgThr	1638
QY	1259	AlaTyrLysArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMet	1278	Db	5404	AGCAGATATGACTCTCTCTCAGGTACACAGCAGCAGCCAGCAGCTCCGCTCCGGG	5463
Db	4324	GCTTACAGCGCAAGTCTAGGGAATGACCTCACACTCAAGCGCTCCAAATGCAATG	4383	QY	1639	ProMetIleThrProAspLeuGluSerGlyValLysMetTyrHisLeuValLysAsnHis	1658
QY	1279	AspAsnLeuGluSerArgValAlaLeuLeuLysGlyLysGluAlaPheAlaGluLeuThr	1298	Db	5464	CCCATGATCACCCAGAGCTTGGAGAGCGGTCTCAAGGTTTGGCATCTGGTGAAGATCA	5523
Db	4384	GACAACCTGAGTCCAGGTGGACCTGGAGTGAAGGAAGCTTTGCGGAGCTTCAGACA	4443	QY	1659	GluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLeu	1678
QY	1299	AspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspTyrArg	1318	Db	5524	GACCATGGTGACAGAGAGGGTGACCGGGCACCACCAAAATGCTGCTGAGATCTACT	5583
Db	4444	GACATCAATGAGCTAACCACTGACTTGGATCGATCGAATCTCTTACCTGGACTACCG	4503	QY	1679	ThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspLeuPheGluThr	1698
QY	1319	ThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLeu	1338	Db	5584	ACCCGGCTTCTAGCCACCAGGGCACCTGCGAGAAATTTGTGGACGACTGTTTGGAG	5643
Db	4504	ACCTATGCCATGAGAGTCTCTTCCAGGCAATTTGAGGACCACTGCTCTGCGGGA	4563	QY	1699	IlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAsp	1718
QY	1339	GluValProGlyTyrArgGlnGlnArgValGluLysGlyLeuLysLeuPheAlaGlnLeu	1358	Db	5644	TTGTTACACATGTCACCGGGGTAGTCTCTCCCTTAGCCATCAAGTACATGTTT	5703
Db	4564	GAGGTACAGGGAATGACAGACAGCAGCTGGAGAAACCCCTGAAACTCTTCGCCAG	4623	QY	1719	PheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrp	1738
QY	1359	IleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPhe	1378	Db	5704	TTCTCTGATGACAGCGCAGACACAGACTATCCAGCAGACAGATGTGGCGCACCTGG	5763
Db	4624	ATCAACAACAAGGTGTTCTTCTGACCTTCATCCGTACACTGGAACCTACAGCGAG	4683	QY	1739	LysSerAsnCysLeuProLeuArgPheTyrValAsnMetIleLysAsnProGlnPheVal	1758
QY	1379	SerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeu	1398	Db	5764	AAAACAACCTGCTTCCACTTCTGTTCTGGGTGAATGCTATCATCAGAACCTCAAT	5823
Db	4684	TCCATGGCAGACCGTGGGAACGTGGCTCTCTCATCATGACAGCGCTTCAGGGT	4743	QY	1759	PheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnThrPhe	1778
QY	1399	GluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGlu	1418	Db	5824	TTTGACATCCCAAGGCGACATCAGATGCTGCTCTCTGTTAGTGGCCAGACCTTT	5883
Db	4744	GAATATGCCACTGATGCTTCAAGCAGCTGCTCTGACCTCATTTGACAAGAACCTG	4803	QY	1779	MetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeu	1798
QY	1419	SerLysAsnHisProLysLeuLeuLeuArgThrGluSerValAlaGluLysMetLeu	1438	Db	5884	ATGGACTCTCTGTTCCACATCAGAGCAGCCAGTAGGCAAGGACTCACCTTCCACA	5943
Db	4804	AACAAGAACCACCCAGCTCTCTCCGAGGACTGAGTGTGTGGCCGAGAGATGCTG	4863	QY	1799	LeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSerAspIle	1818
QY	1439	ThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeu	1458				

Db 5944 CTCTATGCCAAGGATATCCCGCTTATAAGAACTGGTGAAGAACTATGTCAGATATT 6003  
 QY 1819 GlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluGlnSerArg 1838  
 Db 6004 GCCAAGCTCCAGCCATTAGTACCAAGATATGAATGCCCTTACCTCGGAGCAGTCCCGC 6063  
 QY 1839 MetHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGly 1858  
 Db 6064 CTGCATGCTACAGAGTTCATATGCTGTAGCGCCCTCAACGAGATCTACTCATATGTCAGC 6123  
 QY 1859 LysTyrSerGluGluIleLeuGlyProLeuAspHisAspAspGlnCysGlyLysGlnLys 1878  
 Db 6124 AAGTACAGTACGAGCTCATCTGCGGCGACTAGAGCAGGATGAACGCCCGCAGCGCAACGA 6183  
 QY 1879 LeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerLeuAspSerAsnLys 1896  
 Db 6184 CTGGCTACAGGTGGAGCATCTCATCAAGCCATGCTCCATAGAGCTGAAG 6237

## RESULT 4

D86949  
 LOCUS D86949 6730 bp mRNA linear ROD 31-MAR-2000  
 DEFINITION Mus musculus mRNA for plexin 2, complete cds.  
 ACCESSION D86949  
 VERSION D86949.1 GI:1655431  
 KEYWORDS plexin 2.  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

AUTHORS Kameyama,T., Murakami,Y., Suto,F., Kawakami,A., Takagi,S.,  
 Hirata,T. and Fujisawa,H.  
 TITLE Identification of plexin family molecules in mice  
 JOURNAL Biochem. Biophys. Res. Commun. 226 (2), 396-402 (1996)  
 MEDLINE 96400270  
 REFERENCE 2 (sites)  
 AUTHORS Shimizu,M., Murakami,Y., Suto,F. and Fujisawa,H.  
 TITLE Determination of cell adhesion sites of neuropilin-1  
 JOURNAL J. Cell Biol. 148 (6), 1283-1293 (2000)  
 MEDLINE 20191894  
 REFERENCE 3 (bases 1 to 6730)  
 AUTHORS Fujisawa,H.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-AUG-1996) Hajime Fujisawa, Nagoya University Graduate  
 School of Science, Division of Biological Science; Furo-cho,  
 Chikua-Ku, Nagoya, Aichi 464-01, Japan  
 (E-mail:fujisawa@bio.nagoya-u.ac.jp, Tel:052-789-2978,  
 Fax:052-789-2979)

## FEATURES

## source

Location/Qualifiers  
 1..6730  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /tissue\_type="brain"  
 /dev\_stage="E17.5"  
 580..6234  
 /codon\_start=1  
 /product="plexin 2"  
 /protein\_id="BAAL3189.1"  
 /db\_xref="GI:1655432"  
 /translation="MQADNLVSVLLSVAMLLIARGTGMPOYSTFHSNRDFTNHLT  
 VHRRTGAVYVGINRYKITNLTIOVAHKTGEEDNKACYPPLIYOPCEVLTLLNN  
 VNKLIDTSENRLLACGLYQGVCKLLRLDLDLFILVEPSHKREHLYLSVYKGTMYG  
 VIVRSDEGDKLFIQVADGKODYFPTLSRKLPRPESAMLDYELHSDVFSLLKI  
 PSDTLAVSHDFIYIYGAGGVFLVQVETPDGMAINSADLFYTSRIVRLCKO  
 DKPHYSVSLPGCTRAGVEYRLQAYLAKPGEALAQAFINSDEDLVFAIFSKGKQ  
 QVHPDDSAALCAPIRALNQLKRLQCYHGEGLNEUMLWLLGDKVQCTKAPVPIID  
 NFGCLDINGPLGGSTVEGLTYTTSRDLTSVSYVNGYSRVFVGTSGKRLKTKRI  
 -DGPVHGQVGYEMVSVFEDGSPILRDMFSTINQLYVMSRQVTRVPSVCEQYTCG  
 ECLSSDPHCNLCALNMCSRDKORAWENFAAISOCMSLEYHPNSISYSDHSLR  
 LLSLVNDAPNLGETACAGNLTEVEGOVSGGVICISPGKDPVPIPLDQDFGLE  
 LQLRSKTKGKIFVSTEFKYNCSAHQICLSVNSAFRCWCKYRNLCIHTDPTCSQFE  
 GRINVEDCPQLVPTPEILLIPGVEKPIITLKARNLPQPSGQRGYECVLISIQGAVHRV  
 PALRNFSSSVQCNSSQYDGMDISNLAVDFVYVWNGNFIIDNPQDLKVLHYKCAQR

## CDS

ESGCLKADHKFECCGSGERRCTLHQHCPSTSSPWLDSHNVKCSNPQITELTV  
 SGPEGTRVTIHGVNLCIDSEIAHHVOVAGVPCPTPIPEVIAEQTVCMEGHAVIG  
 TTSPGRLCIGECKPEEMTKSHOQYTFVNPVSLSPGIRGSGCTMTITCHYLGAG  
 SSVAYLGNQTCFYGSRMNEIYCVSPSSNGLGPVPSVSDRVRVSSLOFEDIDD  
 PVORIPEWISITSGHTPLTITGNDVIOQEPVRVRFKNGKESVNVCTVNTTUTCL  
 APSITSYRPGLOTVRDEFGFLNNVOSLLIYNDTKFIYYPNPFELLSPGTLQDQ  
 KPGSPITLKGKNCPPASGAKLNTVMIGETPTCTVSETQQLCEPNLLTGQHKVMY  
 HVGWVFSPGVSIVSDLLTLPALISIAAGSLLIIVITVLIAYKKRSRNDLTLK  
 RLOMOMNLESRALECKEAEALQTDINELTSLDRSGIPYLDYRTYAMRVLPFGIE  
 DHPVLEVOGNGQOHVEKALKFLQAOLNNKVFELLTPIRLEORSFMRDRGNVAS  
 LIMTGLQRLLEYATDVLKLDLIDKNLKNHPKLLURTESVSRMLTNWFAFL  
 HKFKECAGEPEFLYCAIQOMEKGPIDAITGEARYSLEDKLLRQOITEYKTLINC  
 VNPONESPEIPVKVNLCDITTVQEKILDVAYKNVYSORPRAVMDLEWROGRIAR  
 VLQODEDITTKIEGDKRLNTLMHYOVSDRVVALVPKQTSYNIPASASISRTSISR  
 YDSSFRYTGSPDSLSRVPMLTDPLESGKVWHLVKNDHGDQKESGDRGKSWSEIYL  
 TRLLATKGTLOKFVDDLLFETLFSVHRGSALPLAKYMPDFELDEQADRHSLHDTVRH  
 TWKSNCLIPRFWVWIRKPOFVFIHKGSIIDACLVSVAQTFMDSCTSEHRLGKSDP  
 SNKLIYAKDIPSYENWERYADIAPKPAISDQDMNAYLAEOFSHATEFNLMSALNE  
 IYSYVSKYSEELIGALEQDEQARQRLAYKVEHLINANSIES"  
 BASE COUNT 1656 a 1850 c 1740 g 1482 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. NO.: 0 Length: 6730  
 Score: 6765.00 Matches: 1268  
 Percent Similarity: 80.24% Conservative: 239  
 Best Local Similarity: 67.52% Mismatches: 357  
 Query Match: 67.72% Indels: 14  
 DB: 10 Gaps: 8  
 US-09-964-956-13 (1-1896) x D86949 (1-6730)  
 QY 25 LeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPheArg 44  
 Db 628 CTGCTAGCTCGTGGGACCACAGGTATGCTCAG-----TACAGCACTTTCCAC 675  
 QY 45 GlyGluProAlaGlu---GlyPheAsnHisLeuValAlaAspGluArgThrGlyHisIle 63  
 Db 676 TCTCAGAATCTGACTGGACTTTCACCATTTGACTGTACACCGAAGAACAGGGCTG 735  
 QY 64 TyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThr 83  
 Db 736 TATCTGGGGGTATCAATCGTCTACAGTTGACTGGCAACCTCACCATCCAGGTGGCT 795  
 QY 84 HisGluThrGlyProAspGluAspProLysCysTyrProProArgIleValGlnThr 103  
 Db 796 CACAAGACAGGGCCAGAGACAGACAAAGGCTTCTACCCACCCTCATTTACAGCCC 855  
 QY 104 CysAsnGluProLeuThrThrThrAsnValAsnLysMetLeuLeuIleAspTyrLys 123  
 Db 856 TGCAGTGAAGTGTATACACTCACCACAATGTCAACAACACTACTCATCTACCTACTCT 915  
 QY 124 GluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeuLeuArgLeu 143  
 Db 516 GAGAATCGCTGCTGGCGTGGCGAAGCCTTACCAGGCTGTTCGAAGCTCCTCGGACTA 975  
 QY 144 GluAspLeuPheLysLeuGlyProTyrHisLysLysGluHisTyrLeuSerGlyVal 163  
 Db 976 GATCACCTCTTCATCCTGCTGGTGGAGCCATCCACAAGAGAACACTACTTGTCCAGTGC 1035  
 QY 164 AsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAspLysLeu 183  
 Db 1036 AATAAGACAGCACCATCATGTGTTGATGTTGCGCTCTGAGGGGGAAGCATGGCAAGCTT 1095  
 QY 184 PheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSerArgLys 203  
 Db 1096 TTTATCGGCACGTGCTGGATGGCAAGCAGGATTACTTCCCTACTCTGTGTCCAGCGCAAG 1155  
 QY 204 LeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheVal 223  
 Db 1156 CTGCCCCGTGACCTGTGCTTTACAGCAATGTGGACTATGACCTCCACAGTGAATTTGTC 1215  
 QY 224 AlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAspIleTyr 243

```
Db 1216 TCCTCCCTCATCAGATTCCTCTGACACCCCTAGCCCTGGTCTCTCATTCTGACATCTTC 1275
QY 244 TyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMet 263
Db 1276 TACATCTATGCTTGGCAGTGGGGTCTTCTACTTCTCACTTCTGACCTGTCACGCCAGAG--- 1332
QY 264 ValSerProGly-----SerThrThrLysGluGlnValTyrThrSerLysLeu 280
Db 1333 ---ACCCCTCAGCGCATGGCCATCAATTCACTGAGACCTCTCTATACCTCAGAAAT 1389
QY 281 ValArgLeuCysLysGlyAspThrAlaPheAsnSerTyrValGluValProIleLeuCys 300
Db 1390 GTGCTCTCTGCAAGGATGACCCCAAGTTCACCTCCTATGTCCTGCTCCCTTTGGCTGC 1449
QY 301 GluArgSerGlyValGluTyrArgLeuLeuGlnAlaIatyrLeuSerLysAlaGlyAla 320
Db 1450 ACAGTGTGGGTGGAATATCGCTTCTGAGCGAGCTTACCTTGCAGGAGCCAGGGAA 1509
QY 321 ValLeuGlyArgThrLeuGlyValHisProAspAspLeuPheThrValPheSer 340
Db 1510 GCTCTAGCTCAGGCTTCAACATCAGCAGCAGGAGATGCTCTGTTGCCATCTTTTCC 1569
QY 341 LysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeu 360
Db 1570 AAGGGCAGAGCAGTACCACCCCTGATGACTCTGCCCTCTGCTGCTTCCCATC 1629
QY 361 LysGlnIleAsnAspArgLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThr 380
Db 1630 CGGGCCATCAACTTGCAGGAGGAGGTTGAGTCTGCTACACCGGAGGAGGCAAC 1689
QY 381 LeuAspLeuAlaTrpLeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIle 400
Db 1690 TTGGAGCTCACTGCTGCTGGGAAGGATGTGCAGTCACCAAGGCGCTGTGCCAATC 1749
QY 401 AspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArg 420
Db 1750 GATGATAACTTCTGGGCTGGACATCAACAGCCTCTGGGAGGCTCCACTCTGTGGAG 1809
QY 421 GlyIleProValPheThrGluAspArgAspArgMetThrSerValIleAlaIatyrValTyr 440
Db 1810 GGACTGACCTGTATACACAGCAGGAGCGCTGACCTCTGCTGCTGCTGCTGCTGCTGCTG 1869
QY 441 LysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLysLysIleArgVal 460
Db 1870 AATGGCTACAGTGTGTTTGTGGGACTAAGAGTGGCAAGCTGAAGAAGATTCGAGCT 1929
QY 461 AspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGlnValVal---AspProGly 479
Db 1930 GATGGTCCCCCATGCTGGGTCCAGTATGAGATGTCTCTGTTCAAGATGGGAGC 1989
QY 480 ProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGlu 499
Db 1990 CCAATCTCCGGGACATGGCTTCTCCATCAATCAGCTATACCTATATGTCATGCTGAG 2049
QY 500 ArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrGlnSerCysGlyGluCys 519
Db 2050 AGACAGGTCAACAGGTGCTGTGTAATCATGTGAACAGTATACACAGTATACACCTGTGAGAGTGT 2109
QY 520 LeuGlySerGlyAspProHisCysGlyTrpCysValLeuHisAsnThrCysThrArgLys 539
Db 2110 CTAAGCTCAGGGATCTCACTGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2169
QY 540 GluArgCysGluArgSerLysGluProArgArgPheAlaSerGluMetLysGlnCysVal 559
Db 2170 GACAAATGCCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2229
QY 560 ArgLeuThrValHisProAsnAsnIleSerValSerValSerValSerValSerValSerVal 579
Db 2230 AGCCTTGAGGTACACCCACAGCATCTCTGTGTCAGATCAGACCCGCTGCTCAGCCTG 2289
QY 580 GluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSer 599
```





Qy	66	GlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThrHisGlu	85
Db	285	GGAGCCATCAACAGGATTTACAAACTGCTAACAAATTGACTCTCTTGGCACCCACGTA	344
Qy	86	ThrGlyProAspGluAspAsnProLysCysTyrProProArgIleValGlnThrCysAsn	105
Db	345	ACAGGTCCGGTAGAGCAATGAAAGTGTATCCCTCCGCCTAGTGTCCAGTCTGC	404
Qy	106	GluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuLeuLeuLeuLeuLeu	125
Db	405	CATGGACTTATCACCAGCAATATGTGAACAAGCTGCTACTCATTTACTTACTCTGATAAC	464
Qy	126	ArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeuLeuArgLeuGluAsp	145
Db	465	AGATTATAGCATGTGGAGCGCTCACAGGGATCTGCCAGTCTCTACGCCCTTGATGAT	524
Qy	146	LeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeuSerGlyValAsnGlu	165
Db	525	CTCTCAAGTTGGTGAGCTCCACCGGAAGAACATTACCTTCTTACTGTCAATGAG	584
Qy	166	SerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAspLysLeuPheIle	185
Db	585	TCTGGGACCATGCTGGTGTTCATCATGTGAGGTTCCAAATGCCCAATAAATCTTTGTT	644
Qy	186	AlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSerArgLysLeuThr	205
Db	645	GGGACTCTTATGTGAAAAATCGAATATTTCCCAACCTCTCCAGCGCAAACTACTG	704
Qy	206	LysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheValAlaSer	225
Db	705	GGCAATGAAGAAAATGCTGAATATGTTGGCTTTGTCTATCAGGATGAATTTGTCTTCA	764
Qy	226	MetIleLysIleProSerAspThrPheThrIleIleProAspPheAspIleTyrTyrVal	245
Db	765	CAACTCAAGATCCCTCAGATACCTTGTCAAAATTCCTACATTTGATATCTACTAGCTG	824
Qy	246	TyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGln-----ProGluMet	263
Db	825	TACAGTTTCAGCAGTGAGCAGTGTGTGTACTACTTAACGCTTCAGCTGGACACCCAGTTA	884
Qy	264	ValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuValArgLeu	283
Db	885	ACATCAGCT-----GACTCAACAGGGGAGCAGTTTTTACGTCCCAAGATTTGCCGCTT	938
Qy	284	CysLysLysAspThrAlaPheAsnSerTyrValGluValProIleGlyCysGluArgSer	303
Db	939	TGCTAGACCATCCCAAGTTCTACTCTTACGTAGAGTTTCCATTTGTTGCATGAAGGAT	998
Qy	304	GlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaValLeuGly	323
Db	999	GGGTGGAATATAGGCTGATACAGGATGCCTATCTGAGCAAGCCAGGAAACGCTTGCT	1058
Qy	324	ArgThrLeuGlyValHisProAspAspAspLeuLeuPheThrValPheSerLysGlyGln	343
Db	1059	AAGAGCTGGGCATCTCTGAAAGGGAAGACATTTCTTTCAGTCTTCTCACAGGACAG	1118
Qy	344	LysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysGlnIle	363
Db	1119	AAGACCGGATCAACCCCCCAAGAGTCTGCTCTGTGGCTCTTTCACCTTGAAGAAGATC	1178
Qy	364	AsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThrLeuAspLeu	383
Db	1179	AAGATAAAATTAAGAGCGCATTCAGTCTTGCTACAGAGCGGATGGAAATCTCTCTG	1238
Qy	384	AlaTrpLeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIleAspAspAsn	403
Db	1239	CCTTGGCTTTAAACAAGAGCTGGGATGATCAATTCACCTCTCTGCAAAATTCAGACAAT	1298
Qy	404	PheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArgGlyIlePro	423
Db	1299	TTCTGTGGCAGGATTTTAACAGCCCTTGGGGGGACATCTACCATAGAGGTACACCG	1358
Qy	424	ValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrValTyrLysAsnHis	443

Db	2439	TTATCTGTGGTCTGGAATGGGCACCTTTGTCATAGACAAACCCCTCAAAACATTCAAGCTCAT	2498	Qy	1156	LysProGlyThrProIleLeuLysGlyLysAsnLeuLeuProProValAlaGlyGly	1175
Qy	801	LeuTyrLysCysGlyAlaMetArgSerCysGlyLeuLysLeuLysAlaAspProAsp	820	Db	3570	AAAGCCAGCTCTCTTATATCAAGGGTCGGAATCTCATT--CCAGCAGCCCGCAGG	3626
Db	2499	CTGTACAAGTCTCGGCCCTAAGGAAAGCTCGCGACTTTGCCTTTAAATCTGACCGTCGA	2558	Qy	1176	AsnValLysLeuAsnTyrThrValLeuValGlyGlyLysProCysThrValThrValSer	1195
Qy	821	PheAlaCysGlyTyrCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysProAla	840	Db	3627	AACTTCAGGCTTAATATACAGTGTGATGTGTGACACTCCGTGTGCCCTTACTGTCTCC	3686
Db	2559	TTTGAGTGTGGCTGTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	2618	Qy	1196	AspValGlnLeuLeuLysGlnSerProAsnLeuLeuGlyArgHisLysValMetAlaArg	1215
Qy	841	GlnGluSerGlnTyrLeuLeuLeuSerGlyAlaLysSerLysCysThrAsnProArgIle	860	Db	3687	GAGACCCAGCTCTGTGGAGTCTCCGAACCTTACTGGAGCAGCATAAGTTACTATTAAA	3746
Db	2619	CTTGAACCCCTTGGATCATGCCAGTACCGCCCAATAGTCGTGCTGACCTGACCCAAAGATC	2678	Qy	1216	ValGlyGlyMetGluTyrSerProGlyMetValTyrIleAlaProAspSerProLeuSer	1235
Qy	861	ThrGluIleLeuProValThrGlyProArgGluGlyGlyThrLysValThrIleArgGly	880	Db	3747	CGCGGAGGGTTTGAGTACTCACCGGCACACTGCAGATTTACTCGGACAGCCTCTGACC	3806
Db	2679	ACCAAGCTGTTCAGACAGACGGCCCGAGGAGGTGGAACCGACTGACCATCACAGGA	2738	Qy	1236	LeuProAlaIleValSerIleAlaValAlaGlyLeuLeuLeuIlePheIleValAla	1255
Qy	881	GluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyValGlu	900	Db	3807	TTGCCTGCCATTATTTGGCATTTGGAGGAGGGGGTTTGTCTGTGTCTATTAATCATC	3866
Db	2739	GAATATCTGGCTTGAGGTTTCAGGACATTCCTTTGGAGTGGGGTGGCGACGTCATG	2798	Qy	1256	ValLeuIleAlaTyrLysArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGln	1275
Qy	901	CysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMetGlyGlu	920	Db	3867	GTGCTCATTTGCCTACAGAGAGAAATCGCGGATCGCCAGCCACCCCTGAAACGACTGCAG	3926
Db	2799	TGTGTCCCTGTGAGAGGCGAGTATATCAGCGCAGAGCAGATCGTTGTGAGATTAAACGAT	2858	Qy	1276	MetGlnMetAspAsnLeuGluSerArgValAlaLeuGluCysLysGluAlaPheAlaGlu	1295
Qy	921	Ala---LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgPro	939	Db	3927	CTGCAATGACAAATCTGGAGTCCCGTGTGGCACTGGAGTGCAAGAGCCCTTTCGACAG	3986
Db	2859	CGCGGAAGGACCCGTGTCACGAGGCTCAAGTGGAGGTCTGTGTAAAGACTGCTCTCAA	2918	Qy	1296	LeuGlnThrAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeu	1315
Qy	940	GluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeu	959	Db	3987	CTCAGACTGACATCCAGAGCTGACCAACGACCTTGACGGGGTGGGAATTCCTTTCCTG	4046
Db	2919	GACTACAGGCTATTCTCCAAAGTCATTACCTTTGTGTGTACCTCTTAAACCGGGTC	2978	Qy	1316	AspTyrArgThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLeu	1335
Qy	960	LysProSerArgGlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeu	979	Db	4047	GAGTACCGTACTTACGGCATCGAGTCTCTTCCCGGGATAGAGATCATCTCTGCTCG	4106
Db	2979	ACACCTTCCGAGGTCCTTTGCTGTGTGGCACCCTGATATCCATTGAAGAAACTATCIG	3038	Qy	1336	ArgAspLeuGluValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPhe	1355
Qy	980	AsnAlaGlySerAsnValValMetPheGlyLysGlnProCysLeuPheHisArgArg	999	Db	4107	AAAGAAATGGAGGTA-----CAAGCCACGTGGAGAAATCCCTGACGCTTTT	4154
Db	3039	AAAGCTGGGAGTCATGCTTGTGTGCTATTGGGGGACGTCCTGTCATGTTTTCATGGCGA	3098	Qy	1356	AlaGlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGln	1375
Qy	1000	SerProSerTyrIleValCysAsnThrThrSerSerAspGluValLeuGluMetLysVal	1019	Db	4155	GGGCAGCTCTCTACCAAGAGCACTCTCTGTGTGACCTTCAITCGGACGCTGGAGCCAG	4214
Db	3099	ACTGCGAAAGAGATCCGCTGTAAGACTCCGACGGGACCAAGCAGGGTAAAGCCGGAATA	3158	Qy	1376	ArgSerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGln	1395
Qy	1020	SerValGlnValAspArgAlaLysIleHisGlnAspLeuVal---PheGlnTyrValGlu	1038	Db	4215	AGGAGCTTCTCCATGAGAGACAGGGGTAACTGCTTCCGCTTCATCATCATGACGCGCTGCAG	4274
Db	3159	CAATCTCTCATCAACCGCCCGCCAGATGAACACTCTGAAGTACATATACACCGAA	3218	Qy	1396	SerLysLeuGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuLeuAspLys	1415
Qy	1039	AspProThrIleValArgIleGluProGluTyrSerIleValSerGlyAsnThrProIle	1058	Db	4275	GGGAGATGGATACGCCACCGCGCTTCTCAAGCAGCTTCTCTCCGACCTCATTTGAGAAG	4334
Db	3219	GATCCAAACCGTCAAAAGATAGAACACAGAGTGGAGCATCGCCAGTGGGGGAACCCCTCTG	3278	Qy	1416	AsnLeuGluSerLysAsnHisProLysLeuLeuLeuArgArgThrGluSerValAlaGlu	1435
Qy	1059	AlaValTyrPheThrHisLeuAspLeuIleGlnAspProGlnIleArgAlaLysHisGly	1078	Db	4335	AATCTTGAAGCAAAACCAACCCCTAAGCTGTGCTGAGGAGGACAGAGTCTGTGGCTGAG	4394
Db	3279	ATAGTAACTGGGATGATCTGGCACTATTAAAGAACCCAAATAAGACTAATAACGGA	3338	Qy	1436	LysMetLeuThrAsnTyrPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGly	1455
Qy	1079	GlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAla	1098	Db	4395	AAGATGCTGACAACTGGTTTACGTTCTCTCTACAGTCTCTTAAAGAAATGCTGCTGG	4454
Db	3339	GATGTGGAGAAAGAAATACTGCATCTTTACAATGACACCAACCATGTTGTGTTGGCA	3398	Qy	1456	GluProLeuPheSerLeuPheCysAlaIleLysGlnGlnMetGluLysGlyProIleAsp	1475
Qy	1099	ProAlaLeu-----AlaLeuGlyProAspHisGlnSerAspLeuThrGluArgPro	1115	Db	4455	GAGCGCTCTTCTATCTTGCATCTGCCATCAAGCAACAGATGAGAAAGAGACCTATTGAT	4514
Db	3399	CCCTCCGTTGACAATCCCTTGGAGAGCCCTCCGAGAGATGGGAC-----CGTCCT	3449	Qy	1476	AlaIleThrGlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIle	1495
Qy	1116	GluGluPheGlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsn	1135	Db	4515	GCCATCACTGGAGAGCCGCTACTCACTGAGGAGCAAGCTGATCCGTCACAGATC	4574
Db	3450	GATGAGATAGCTTCATTAATGGACAATGTTTCATGCCCTCTTATTTGTCACACATACCAGC	3509	Qy	1496	AspTyrLysThrLeu-----	1500
Qy	1136	PheThrTyrTyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeu	1155	Db	4575	GACTACAGACACTGAATCCTTGTGCTGATGAGTGTGGCTGAGCGCATGAATCATGTTGC	4634
Db	3510	TTCTTTATACCCGGATCTCTGTTGAAACCGCTGCGGCATCTGGAACACCTTTCAGCTG	3569				





Db		...:     ...:				2113	TCACCTTCTCCCGGGAAGTGCGCCCATCACACAGGGTCAGGAGACCAGCGGGTGGT	2172
Qy	634	GlnLeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerThrSerPheValPhe	::::			634		
Db	2173	AAGCTCTACCTCAAGTCCAAGAGACAGGAAAGATTGCATCTGTGGACTTTGTCTC				2173		2232
Qy	654	TyrAsnCyssSerValHisAsnSerCySLeuSerCysValGluSerProTyArgCysHis				673		
Db	2233	TATAAATGCAAGTGTTCACCAAGCTCTGCCCTGGCATGTGTCAAAGGCCTCTTC				2292		
Qy	674	TrpCysLysTyrArgHisValCysThrHisAsnProLysThrCysSerPheGlnGluGly				693		
Db	2293	TGTTGTCAAAATACCGCCATGTTTCACCAACAATGCGGCTGCAGTGC				2352		
Qy	694	ArgValLysLeuProGluAspCysProGlnLeuLeuArgValAspLysLysLeuValPro				713		
Db	2353	CGAGTCACATGCTGAGGACTGCCACAGATCCTGCCCTTACCCACATCTATGTGCCA				2412		
Qy	714	ValGluValIleLysProIleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGly				733		
Db	2413	GTGGGTGTGTGAAGCCCATCACCTGGTGTGCACGGAACCTGCCGAGCCACAGT				2472		
Qy	734	GlnArgGlyTyrGlucylsIleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeu				753		
Db	2473	CACGAGGGTACAGATGTCCTTTCCATATCCCAGGAGCCCGCCGAGTCACTGC				2532		
Qy	754	ArgPheAsnSerSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGlu				773		
Db	2533	CGCTTCACAGCTCCAGCCCTGCAGTGCAGAACTCTCTGTACTCTATGAGGGAA				2592		
Qy	774	IleAsnAsnLeuProValGluLeuThrValValTrpAsnGlyHisPheAsnIleAsp				793		
Db	2593	GTCAAGCATGCTGCTGTGAACCTGTGCGTTGTGTGGAATGGCAACTTTGTCA				2652		
Qy	794	ProAlaGlnAsnLysValHisLeuTyrLysCysGlyAlaMetArgGluSerCysGlyLeu				813		
Db	2653	CCCCAGAACATCAGAGCCCACTCTCAAGTGGCCAGCCCTTGGCCAGAGATTGT				2712		
Qy	814	CysLeuLysAlaAspProAspPheAlaCysGlyTrpCysGlnGlyProGlyGlnCysThr				833		
Db	2713	TGCGCTCAAGCAGACCCACCGCTTTGAGTGC				2772		
Qy	834	LeuArgGlnHisCysProAlaGlnGlu---	SerGlnTrpLeuGluLeuSerGlyAlaLys	852				
Db	2773	CTTCGTCAACCATGCGCCAGCTGACTCACCAGCCTCTGGATG	CGCCACCATGGCAGC	2832				
Qy	853	SerLysCysThrAsnProArgIleThrGluIlelleproValThrGlyProArgGluGly	872					
Db	2833	AGCGCTGCACAGCCCAAGATTCTCAAGCTCTGCCACAGACAGAGCCCTCGC	AGGGA	2892				
Qy	873	GlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuPheArgAspIleAlaSer	892					
Db	2893	GGGACTCGACTAACCATCACCGGTGAGAATCTGGGCTCTCGGTTGAGGAT	TGGCGGCTA	2952				
Qy	893	HisValLysValAlaGlyValGluCysSerProLeuValAspGlyTyrlleProAlaGlu	912					
Db	2953	GGCGTGCACTGGGCAAGGTGCTTTGCAGCCCTGTGGAGACGAGTACATCAGC	CGAGAA	3012				
Qy	913	GlnIleValCysGluMetGlyGluAlaLysProSerGln--HisAlaGlyPheValGlu	931					
Db	3013	CAGATTCTCTCGAGAGATTGGAGATGCCAGCACGCTCGCGGCCCATGATGCC	CTGTGGAG	3072				
Qy	932	IleCysValAlaValCysArgProGluPheMetAlaArgSerSerGlnLeuTyrTyrPhe	951					
Db	3073	GTGNGTGTGGGACTCTCTCCCTCACTACCGAGCGCTGTCCCTTAACCGCTT	CACTTTT	3132				
Qy	952	MetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMetSerGlyGlyThrGln	971					
Db	3133	GTGACACCGACCTCTACCGTGTGAGCCCTCCCGTGGGCTCTGTCTGAGGTAC	CTGTGG	3192				
Qy	972	ValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnValValMetPheGlyLys	991					



```

/notes="kozak consensus"
185...5800
/genes="sex"
185...5800
/genes="sex"
/codon_start=1
/product="SEX protein"
/protein_id="CAA61132.1"
/db_xref="GI:1200235"
/db_xref="SWISS-PROT:P51805"
/translation="MPSVCLLLLLFLAVGALGNRPFAFVVTDTTLTHLAVHVTGE
VFVGNVNRVFKLAPNLTELRAHVTGPEVDNARCPPESMRVCAHRLAPVDNINKLLLI
DYAARRLVACGSIQWQICFLRLDDLEFKLGEHPHRRKEHYLSGAGPEPDSMAGVYEQGO.
GPKLFTVGTADKSEYFTLSRKLISDESDAMFSLVYODEFVSQIKIPSTFYSLV
YPALDIYIYGFVSASFVFLTLQDLTQOQTLDTAGKFFTSKIVRMKAGDSFYSY
EFTFGCSWRVSVLVSUHAHLKPLGALLAQLGVPADDEVLFTIFSOGOKNRASPPRO
TEILCFTLSNINAHIRRIOSCYRGECTLALPWLNLKELPCINTPXOINGNFCGLVLN
QPGLGLHVIIEGLPLADSTGMSAAYTYRQHSVFIETGRSGSLKRVYDGFODAHLL
YEIVPVVDSPIILRDLLFSPDRHRIYLLSEKQVSQLPVECTEQYOSCAACLGSDDPHC
GWCVLRRHRCREGACLGASAPGHFAELSKCVQVRVPPNNVSVTSPGVLTVTLHNVP
DLGAGVSCAFEAANEAVLLPSELCPSPSLQELRALTRGHGATRTVRLQLLSKET
GVRFAGADVFVYNCVLOSMSCVSPYPCWCKYRHTCTSRPHECSFEQGRVHSPEG
CPELLPSGDLILPVGMOPILTRAKNLPOQSGOKNYECVVRQGRQORQPAVFNSS
SVQCQNASYSYEGDEHGDTELDSVYWDGDFPIDKPSFRALLTKNAQRPSCGLCLK
ADPFNGCWCISHRCOLRTHCPAPKTNMHLISQKTRCSHRPRITQIHLVPGKEGT
RVTVIGENGLLSREVLVAGVRCNSIPAEYISAERIVCEMEESLVPSPPPGVELC
SECOFVRRDAKATVCSPLTGLPSQAPITLAIADRANISSPGLIYTVTDPTVTRLEP
TWSLNGSTAITVSGTHLLTVQBPVRVRAKYRIETNTTCOVINDTAMLCRKPGLFLGR
POPRAGEHDEGEFLDHVQTARSLNRSSEFTYYPDPSFELPGSGVLVDYKPGSHVIL
KGNLIPAAAGSRNLITVLGGPQSLTSDTLCCDSFSQTRGQRPVWLVLGGLEFV
LGLTHLSAERALLTPAMMGLAAGGGLLLLAITAVLVAIKRKTODADRLKRLQMDN
LESRALECKEAEFAELQDINELTNHMDVQIPELDYRTAVRVLPFGIEAHPVLKEL
DTPPNVKALRLFGQLLSRAFLVTFHTLEAQSFSMRDRGTVASLTVALQSLRDY
ATGLLKOLLALTEKLENSKNHPKLLRLTAEKMLTNWFTLLHKKFLKECAQELP
FLYLCATKOQMEKPIDAITGEARYSLSEDKLIRQOJDYKTLTLHLCVPENEGSAQVP
KVLNCDSTQAKDKLLDVTYKGPISQRPKAEDMDLEWROGRMIRILQDEDTVTKI
ECMDKRLUNSLAQVTDGSLVALPVQVSAYNMANSFTFRSLRSYLSLRTASSPDS
LRSPAMITPDQGTGLKHLVKNHHDHREGRDKRGMVSEIYTLRLATKATGLQKF
VDLFEFTVFAHGRSALPLAKYMFDFLDEQADQROISDPDRHRTWKNCLPLRFVW
NVTKNPOFVDFIHNKSTIDACLKYVOTFMDSCSTSEHRLGKDSNKLKLVAKDIPNY
KSWERYRYRDIAMKASISDQMDAYLVEQSLHASFVLSALNELFYVTKYRQEIL
TALDRASCRRKHLRQLEQIILISLVSDS"
BASE COUNT 1116 a 1985 c 1826 g 1110 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 6039
Score: 5991.00 Matches: 1136
Percent Similarity: 75.34% Conservative: 294
Best Local Similarity: 59.85% Mismatches: 440
Query Match: 59.97% Indels: 28
DB: 9 Gaps: 14

US-09-964-956-13 (1-1896) x HSSEXGENE (1-6039)
QY 1 MethlysAlaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValcglyMet 20
   ::: |||||
Db 176 CTGCGGCCCATGCCCC-----TCGTGCTGCCTCCTCCTG---CTGCTCTTCTTCCGCGTG 226

QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40
   ||| |||
Db 227 GGGGGGGCC-----CTGGGCAACAGGCCCTTC-----CGTGCCTTC 262

QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60
   ||| |||
Db 263 GTGGTG-----ACAGACACACGCGTTTACCACCTGGCTGTGTCACCGGCTGACT 310

QY 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80
   ||| |||
Db 311 GGGGAGGTGTCGNGGGCGCAGTCAACCGAGTCTTTAAGCTGGCGCCCACTGACTAC 370

QY 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100

```

D	b	1451	CGCCAGCACTCTGTGGTCTTCATTTGGCACGCCGACGGCAGCTTGAAAGAAGGTCCGGGTC	1511
Q	y	461	AspGlyProArgGlyAsnAlaLeuGlnTyThrValGlnValValAspProGlyPro	480
D	b	1511	GATGGCTTCCAG--GATGCCACCCTTATGCACACAGTCCCCTGGTGGATGCGACGCC	1567
Q	y	481	ValLeuArgAspMetAlapheSerLysAspHisGluGlnLeuTyrlleMetSerGluArg	500
D	b	1568	ATCCTCGGACACTGCTTCTTCAGCCCGGACACCGGCACATCATCTATCTCTAGTAGAAG	1627
Q	y	501	GlnLeuThrArgValProValGluSerCysGlyGlnTyrglnSerCysGlyGluCysLeu	520
D	b	1628	CAGGTGACCGAGCTCCCGTGGAGACTGTGCAGATGACCGAGCTGCGAGCGCTGCCTG	1687
Q	y	521	GlySerGlyAspProHisCysGlyTrpCysValLeuHisAsnThrCystThrArgLysGlu	540
D	b	1688	GGTCCGGGACCGCAGCTGCTGGTGTGTGTCGCGACACAGGTGCTCCGCCGAAGGG	1747
Q	y	541	ArgCysGluArgSerLysGluProArgArgPheAlaSerGluMetLysGlnCysValArg	560
D	b	1748	GCCTGTCTGGCGCCTTCGCCCACACGGCTTCTGAGGAGCTGAGCAAGTGTGCCAG	1807
Q	y	561	LeuThrValHisProAsnAsnIleSerValSerGlnTyrrAsnValLeuLeuValLeuGlu	580
D	b	1808	GTGCGGCTCGGCCAACAAATGTGCAGTCACGTCCACTCGGGTGCAGCTGACCGTCACC	1867
Q	y	581	ThrTyrrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSerGlu	600
D	b	1868	CTCACAACGTGGCAGACCTTCAGTGGGGCGGTGAGCTGCCTTCGAGGCGCGCGGGAG	1927
Q	y	601	MetAspGly---LeuValValGlyAsnGlnIleGlnCysTyrrSerProAlaAlaLysGlu	619
D	b	1928	AACGAGCGGCTCTGTGCTGCCCTCCGGTGAAGTCTCTGCCCTCACCTCCCTCCAGGAG	1987
Q	y	620	ValProArgllelleThrGluAsnGlyAspHisHisValGlnLeuGlnLeuLysSer	639
D	b	1988	CTCCGAGCTCTTACCAGGGGCGATGGGGCCACCCGACACTGTGGGGTGCAGCTTCTCTCC	2047
Q	y	640	LysGlnThrGlyMetThrPheAlaSer-Thr-Ser-PheValPheTyrrAsnCysSerValHis	659
D	b	2048	AAGAGACAGCGGTGAGTTTGGCGGTCTGACTTTGTCTCTACAACTGCAGCGTCTC	2107
Q	y	660	AsnSerCysLeuSerCysValGluSerProTyrrArgCysHisTrpCysLysTyrrArgHis	679
D	b	2108	CAGTCGTGCATGTCTCTGTGGCAGCCCTTACCCCTGCCACTGGTGTAAAGTACCGGCAC	2167
Q	y	680	ValCysThrHisAspProLysThrcysSerPheGlnGluGlyArgValLysLeuProGlu	699
D	b	2168	ACGTGTACCAGCGCCCCCAGAGTGTCTCTCCAGGAGGCGAGGTCCACAGCCCTGAG	2227
Q	y	700	AspCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysPro	719
D	b	2228	GGCTGCTCTGAGAPCTTCGCCAGCTGGGAGACTCTCTGATCCCGCTTGGGGTCATGCAGCCT	2287
Q	y	720	IleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrrGluCys	739
D	b	2288	CTTACCTTGGGGCTAGAACTTACCTCAGCGGAGTCCGGGCCAAGAAGAACTATGAGTGC	2347
Q	y	740	IleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArgPheAsnSerSerSer	759
D	b	2348	GTGTCGGGTGCAGGGGCGGCACACGCGGTGCTCCCGTGGCGCTTCAACAGCAGCAGT	2407
Q	y	760	ValGlnCysGlnAsnThrSerTyrrSerTyrrGluGlyMetGluIleAsnAsnLeuProVal	779
D	b	2408	GTGCAGTGCAGAAACGCTCGTACTCTCTATGAAGGTGATGAGCATGTGTGACACCGAGCTG	2467
Q	y	780	GluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsnProAlaGlnAsnLysVal	799
D	b	2468	GACTTCTCCGTGGTCTGGGATGGAGACTTCCCCATAGCAACAGCTCCCGACTCCCGAGCC	2527
Q	y	800	HisLeuTyrrLysCysGlyAlaMetArgGlnSerCysGlyLeuCysLeuLysAlaAspPro	819
D	b	2528	CTCCTGTACAAGTGTCTGGGCGACGGGCCAGCTGTGGCTCTGCCCTCAAGCTGATGCC	2587

Qy	820	AspPheAlaCysGlyTrpCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysPro	839
Db	2588	CGGTTCAACTGTGGCTGGTGCATCTCAGACGACAGGTGCCAGGTCCGGAGCCACCGTGC	2647
Qy	840	AlaGlnGluSerGlnTrpLeuGluLeuSerGlyAlaLysSerLysCysThrAsnProArg	859
Db	2648	GCCCGAAGACCAACTGGATGCACCTCAGCCAGAGGGGCACCCGGTGCAGCCACCCCGC	2707
Qy	860	IleThrGluIleIleProValThrGlyProArgGluGlyGlyThrLysValThrIleArg	879
Db	2708	ATCAGCGAGATCCACCCTCTCGTGGGGCCCAAGGAAGAGGACCCCGGGTCACCATCGT	2767
Qy	880	GlyGluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyVal	899
Db	2768	GGTGAGAACCTGGGCGCTTGTCCCGAGAGGTGGGC-----CTGCGGGTGGTGGCGGT	2821
Qy	900	GluCysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMetGly	919
Db	2822	CGTTGCAACTCCATTCCGGCCGAGTACATCAGTCTCAGAGGATCGTGTGAGATGGAG	2881
Qy	920	GluAla---LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArg	938
Db	2882	GAGTCGTGTGTGCCAGCCGCGCCCGGGGCCGCGTGGAGCTGTGTGTGGTGAGCTGTCA	2941
Qy	939	ProGluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAsp	958
Db	2942	GCCGACITTCGCGACCGAGTCTGGAGCAGGTCTACAGCTTTGTGACCCCAACGTTTGACCA	3001
Qy	959	LeuLysProSerArgGlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsn	978
Db	3002	GTGAGTCCCAGCGGTGCGCGCGCGTCCGGGGGCACACGGCTTACCATCTCAGGCAGCTCT	3061
Qy	979	LeuAsnAlaGlySerAsnValValMetPheGlyLysGlnProCysLeuPheHisArg	998
Db	3062	CTGATCTGTGCACAGGGTACAGTCACTGTRAGGACACGCGAGTGCCAGTTGTGAAG	3121
Qy	999	ArgSerProSerTyrIleValCys--AsnThrThrSerSerAspGluValLeuGluMet	1017
Db	3122	AGAGTGCCAAAGCGGATCGTGTGATCTCACTCTCCACCTCGGGCCCGCCAGCGCC	3181
Qy	1018	LysValSerValGlnValAspArgAlaLysIleHisGln---AspLeuValPheGlnTyr	1036
Db	3182	CCCATCACATGTGCATGTGACCGGGCTAACATCTCCAGCCCGGGGTCTCATCATCACCTAC	3241
Qy	1037	ValGluAspProThrIleValArgIleGluProGluTrpSerIleValSerGlyAsnThr	1056
Db	3242	ACTGAGAACCCACCGTCAACCGCTTGAGCCCACTGGAGCATCATCAATGGAAGCACT	3301
Qy	1057	ProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLys	1076
Db	3302	GCCATCATCTGTGAGTGGGACCACTGTGACGCTCAGGAGCGCCCGGGTCCGTGCCAAG	3361
Qy	1077	HisGlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCys	1096
Db	3362	TACCGGCGCATTGAGACCAACCAATACATGCCAAGTGATCAACGACGACTGCCATCGTGT	3421
Qy	1097	GlnAlaProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgProGlu	1116
Db	3422	AAGCCCCCGGCATCTTTCTTGGCGGCCCGCCAGCCTCGGGCGCAAGCGACACCTGAT	3481
Qy	1117	GluPheGlyPheIleLeuAspAsnValGlnSerLeuIleLeuAsnLysThrAsnPhe	1136
Db	3482	GAGTTTGGCTTCTGTGTGACACAGTGCAAACGGCCCGCTCCCTCAACCGCTCTCCCTTT	3541
Qy	1137	ThrTyrTrpProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeuLys	1156
Db	3542	ACCTACTACCTGATCCCGAGCTTTGAGCGCTGGGGCCCTCTGGGTGCTGGAGCTCAAA	3601
Qy	1157	ProGlyThrProIleIleLeuLysGlyLysAsnLeuIleProProValAlaGlyGlyAsn	1176
Db	3602	CCGGGCTCCCGAGTGTGCTGAAGGCAAGAACCTGATTCCCGGGGACCGCGGAGCTCC	3661

QY	1177	ValLysLeuAsnTyrThrValLeuValGlyGluLysProCysThrValThrValSerAsp	1199
DB	3662	---CGCCTCAACTACACTGTGTGATGAGAGGCCGCGTGTTCGGCTCACTGTCTCGGAC	3718
QY	1197	ValGlnLeuLeuCysGluSerProAsnLeuIleGlyArgHisLysValMetAlaArgVal	1216
DB			
DB	3719	ACACAACCTCTGTGGCCACTACCCAGCCAGACTGGCCGCGCAGCCTGTCAATGGTGTGGTG	3778
QY	1217	GlyGlyMetGlnTyrSerProGlyMetValTyrIleAlaProAspSerProLeuSerLeu	1236
DB	3779	GGTGGCCTGGAGTCTGGCTGGGCACCCCTGCACATCTCGCAGAGCGGGCGCTGACCCCTA	3838
QY	1237	ProAlaIleValSerIleAlaValAlaGlyGlyLeuLeuIleIlePheIleValAlaVal	1256
DB	3839	CCGGCCATGATGGGGCTGGCGGGGGGTGGGCTCTGCTGTGTCGTCATCAGACGCGT	3898
QY	1257	LeuIleAlaTyrLysArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMet	1276
DB			
DB	3899	CTGGTGGGTACAGGCGCAAGACTCAGGACCGCGGACCGTACCCTCAAGCGTTCAGAGCTG	3958
QY	1277	GlnMetAspAsnLeuGluSerArgValAlaLeuGluCysLysGluAlaPheAlaGluLeu	1296
DB	3959	CAGATGGACAACCTGGAGTCCCGTGTGGCCTGGAGTGCAAGGAAGCTTTTTCGAGAGCTG	4018
QY	1297	GlnThrAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAsp	1316
DB	4019	CAGACGGACATCAATGAGCTGACTAAACCACATGGACGAGGTGCAGATCCCTTCCTCGGAC	4078
QY	1317	TyrArgThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLeuArg	1336
DB	4079	TACCGGACTTTACCGCTGGCGGTCTTCCCGGGCATCGAGGCCACCGCGTGTCTCAAG	4138
QY	1337	AspLeuGlnValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAla	1356
DB	4139	GAGCTGGATAGCCCA-----CCCAAGCTGGAGAGGCCCTCGCCCTCTTCGGG	4186
QY	1357	GlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArg	1376
DB	4187	CAGCTGTGCACAGCGCGCGTTCGTGCTTACCTTCACACAGCTGGAGGCCACAGC	4246
QY	1377	SerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSer	1396
DB	4247	AGCTTCTCCATGCGGACCGCGCACCGTGGCTCTCCTCACATGTGGCGCTTCAGAGC	4306
QY	1397	LysLeuGlnTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsn	1416
DB	4307	CGCCTCGACTATCCACGGGCTGCTCAAGCACTCTGTCGCCGACCTCATCGAGAGAAC	4366
QY	1417	LeuGluSerLysAsnHisProLysLeuLeuLeuArgArgThrGluSerValAlaGluLys	1436
DB	4367	CTCGAGAGCAAGAACCACCCCAAGCTGCTCTACGCAGGACAGACTCAGTGGCTGAGAAG	4426
QY	1437	MetLeuThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGlu	1456
DB	4427	ATGCTTACCACACTGGTTCAGTTCTGCTGCATTAAGTTTCTGAAGGAGTGTGCTGGGGAG	4486
QY	1457	ProLeuPheSerLeuPheCysAlaIleLysGlnGlnMetGluLysGlyProIleAspAla	1476
DB	4487	CCTCTCTCTGCTTACTGTGCATCAAGCACACATGGAGAGGGCCCCCATGATGCC	4546
QY	1477	IleThrGlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAsp	1496
DB	4547	ATCACGGCGGACGACGATACTCCTTCGAGGAGGACAAAGCTCATCCTCGTCAGCAGATCGAC	4606
QY	1497	TyrLysThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluValPro	1516
DB	4607	TACAAGACACTGACCTTCACTCGCTGTGTCCCGAGAACACGAGGCGCGCCAGGTGCCCA	4666
QY	1517	ValLysIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIle	1536
DB	4667	GTGAGGTTCTCAACTGTGCACATCATCCAGGCCACCAAGCATAAGCTGTCTGGACACTGTG	4726
QY	1537	PheLysAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTrpArg	1556

[illegible]

LOCUS D86950 5982 bp mRNA linear ROD 31-MAR-2000  
DEFINITION Mus musculus mRNA for plexin 3, complete cds.  
ACCESSION D86950  
VERSION D86950.1 GI:1655433  
KEYWORDS plexin 3.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (sites)  
Kameyama,T., Murakami,Y., Suto,F., Kawakami,A., Takagi,S.,  
Hirata,T. and Fujisawa,H.  
Identification of plexin family molecules in mice  
Biochem. Biophys. Res. Commun. 226 (2), 396-402 (1996)  
96400270  
2 (sites)  
Shimizu,M., Murakami,Y., Suto,F. and Fujisawa,H.  
Determination of cell adhesion sites of neuropilin-1  
J. Cell Biol. 148 (6), 1283-1293 (2000)  
20191894  
3 (bases 1 to 5982)  
Fujisawa,H.  
Direct Submission  
Submitted (07-AUG-1996) Hajime Fujisawa, Nagoya University Graduate  
School of Science, Division of Biological Science; Furo-cho,  
Chikua-ku, Nagoya, Aichi 464-01, Japan  
(E-mail: fujisawa@bio.nagoya-u.ac.jp, Tel: 052-789-2978,  
Fax: 052-789-2979)  
Location/Qualifiers  
1. 5982  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="brain"  
/dev\_stage="E17.5"  
144..5762  
/codon\_start=1  
/product="plexin 3"  
/protein\_id="BAAI3190.1"  
/db\_xref="GI:1655434"  
/translation="MPTVCLLPVLTFFITGGCLGSSRPFTFVVTDTTLLHVRVTG  
EVFGVNRVFKPLATTELRAHTPTIEDNARCPYPPSMRVCSRLHLPVDNVLKLL  
IDYARLVACGSGTQIGICQFLRLDLDLFLKGLPEHRKHEHLSGAGEPMSMAGVYEQ  
GPKSLFVGTADGKSEYFTPLSSRLKIDDESGDMFSLVYQDEFVSVSQIKTPSDLS  
LYPAFIYITVFGVSASVYFTLQDQTLTQDLDLAGEKFFTSIVKMGADSEYIS  
VEFTIGCSWRGVNLRHAGLAKPGLLAQALGPADEDLVLTIFSGQKNRANPPR  
QTLILFTLSINAHIRRIQSCYREGTALPWLNLKELPCINTPLQINGFCGLVL  
NQLGGLHVIIEGLPLDADTGMASVAATVYHOSVVFITGRSGNLKVRVDGSDAQ  
LYFTSVVQSGPILDLFSPDRHILYLLSEKVSQLPVETCEOYLSCAACLGSDPH  
CGWVLOHRCREGACPGASAPHAELSKLCIQVRPNVNVSTSSGVLTVARNV  
PDLVSGVSCSEEVTESEAILLPSELKCPSPSLQELQTLINGHATHTVRLQLLSME  
TGVRPAGVDFVYNCALQSCMSCVSPYPCWCKYRHVCTSHPECFSQEGRVHSPE  
GCPETLPQGDLLIPVGMQPLTLRAKNLPQSGQKNEYCVVRQGRHVRPAVRFS  
SSVQONASYEVEGDEFGLTDFSVWMDGFPIDKPPSFALLKWAQRPSCGLCL  
KADRFNGWCISHRCCOLRAHCPAPKSNMHPKOGARCSHPRITQIHLPTGKEGG  
TRVTIGENGLITSREVGIRVAGVRCNSIPTPEYVSAERIVCEMESLVPSPPPGAEL  
CVGCSADFTSQOQLYSEVFTFDVRSFSPGSGGTLRLISGLSDRGSRVTVIR  
DGEQFVRDAEAIVICISPVSTLGPSPITLAIIDANISNTGIYITQDPTVTHLE  
PTWISTINGSTISVGTHTLTPQEPVRKRYGIETNTCQVINDTAMLCAPGIFLG  
HPQRAOCEHDEFGFIDLVHQAAARSLNRSFTYYPDPSPFPLGSGVLDPVPGSHV  
LKGNLIPAAAGSRLNTVLIIGOPCALTVSDTOLLCDSPSQTGRQPMVLVGGLEF  
WGLTHITADRALTLPAWGLAAGGGLLAILTVVLVAYKKRTQADRTLKRLQLOMD  
NLESVALEKAEFAELQDINELTNHMDGVOIPEDYKTYAVRVLFPGIEAHPVLKE  
LDTPNVNLEKAEFGQLHSRAFLFTFTLEAQSFSMRDRGTVASLTMVALQSRLD  
YATGLLKQLLADLIEKNLESKNHPKLLRRTESVAEKMLTNWFTLLHKLFEKAGEP  
FLFLYCALKQOKEGPIDAITCEARYSLSEDKLIRQIDYKTLTLHVCPESEGAQV  
PVKVLNCSITQAKDLDTVYKGIPIYSPORKAEDMDLEWROGRMARILODEDITTK  
LECDKRWKNSLAHQVTDGSLVALVPQVSNMANSFTFESLSRYESLLRAASPD  
SILRSKRWKNTFDQERAGTKLWHLVRNHDHTHREGDRGSKMSEIYTLRLATKGLQK  
FVDDLFTFVSTAHRSALPLAIKYMFDLDAQDQRLSDPDVDRHTKNSCLPLRFW  
VNTKNQCFEVDIHKNSITDACLISVVAQTFMDSCSTSEHRLGKSPSKNLLYAKDIPN  
YKSWERYRIOTAKMASISQDMDAYLVEQSLRHANDFNVLSELSALVELYFYVTKYRQEI  
\*LTSLDASCRKHKLRQKPEQIITLVSSSS"  
1258 a 1704 c 1644 g 1375 t 1 others  
BASE COUNT

ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 5982  
Score: 5984.50 Matches: 5132  
Percent Similarity: 74.97% Conservative: 291  
Best Local Similarity: 59.64% Mismatches: 440  
Query Match: 59.90% Indels: 36  
DB: 10 Gaps: 13  
US-09-964-956-13 (1-1896) x D86950 (1-5982)  
QY 8 TrpThrCysLeuLeuSerHisLeuLeu-----MetValGly-----Met 20  
Db 139 TGGTCATGCCCACTGTCTGC-CYTCCTCCCATTCCTTCCATCAGGAGGGTGCCTG 197  
QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
Db 198 GGTAGCAGC-----AGGCCATTT 215  
QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
Db 216 CGTACCTTTTGGTACACATACCACCT--CTCCTCACCCTGGCTGTCACCGAGTCACT 272  
QY 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
Db 273 GGGGAGGTCTTTAGTGGCGTGAATCGAGTTTCAAGCTGGCTGCTCAAAACCTAACTGAG 332  
QY 81 LeuValThrHisGluThrGlyProAspGluAspProLysCysTyrProProArgIle 100  
Db 333 CTACGGGCCCATGTACAGGGCCCATGTAGGAGCAATGCTCGCTGCTACCCACCTCCTAGC 392  
QY 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIle 120  
Db 393 ATCCGTGTGTCTCCCAACCCCTGGTGCCTGTGGACAATCTGAACAAGCTGCTCCTCATA 452  
QY 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
Db 453 GACTATGCAAGCCCGTCTGTGTAGCTGTGGCAGCATCTGGCAGGCGATCTGCCAGTTC 512  
QY 141 LeuArgLeuGluAspLeuPheLysLeuGlyLupProTyrHisLysLysGluHisTyrLeu 160  
Db 513 CTGCGTCTGGATGACCTCTTCAAGTGGGTAGGCCCCACCATCGCAGGAGGACCTATCTG 572  
QY 161 SerGlyValAsnGlySerGlySerValPheGlyValIleValSerTyrSerAsnLeuAsp 180  
Db 573 TCAGGGGCCAGCAGCGCTGATTCATCGCTGGCGCTCATTTGTTGAACAGGTCCAGGGGCT 632  
QY 181 AspLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSer 200  
Db 633 AGCAAGTTGTTGTGGCACCCTGTGTGTAGTGGCAAGTCTGAGTACTTTCACCTTGAGT 692  
QY 201 SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAsp 220  
Db 693 TCCCGTAAGCTCATCGACGATGAGGACAGTGGAGATATGTTCACTGGTGTACCAAGAT 752  
QY 221 GluPheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPhe 240  
Db 753 GAGTTTGTCTCTCTCAGATCAAGATCCCTTCAGACACATGTTCTGTGTACCTGCTCTT 812  
QY 241 AspIleTyrTyrValTyrPheSerSerGlyAsnPheValTyrPheLeuLeuGln 260  
Db 813 GACATCTACTACATATATGCTTTTGTGTCGCCCTCCCTTCGTGTACTTCTTAACTGTTG 872  
QY 261 ProGluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeu 280  
Db 873 CTGACACCCAGCAGCAGCGCTGTTGGATACAGCAGCGGCAAAATCTTCACATCCAAATC 932  
QY 281 ValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCys 300  
Db 933 GTCCGATGTGTCCAGGGGACTCAGAGTTCTACTTCTATGTATGAGTATCCCTATTGGCTGC 992  
QY 301 GluArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAla 320

Db 993 TCCTGGCGTGGTGGAGTACCGCTGGTGGAGAGTCCCATCTGGCCCAAGCCAGCCGCTG 1052  
Qy 321 ValLeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSer 340  
Db 1053 CTATTGGCCAGGCTCTGGGTGTGCAGCCGAGGAGTGTCTCTTACCAATCTCTCT 1112  
Qy 341 LysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeu 360  
Db 1113 CAGGGCCAGAACAGACAGCCACCCCTCGGCAGACCATCTTTGGCTTTTACCCCTC 1172  
Qy 361 LysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThr 380  
Db 1173 AGCAGCATTAATGCCCCACATCGCGGCCGAATCCAATCGTGTACCGTGAGGGCAGC 1232  
Qy 381 LeuAspLeuAlaTrpLeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIle 400  
Db 1233 CTGGCCCTACCTGGCTACTGAATAGGAGCTGCCCTGCATCAACACCCTCTGCAGATC 1292  
Qy 401 AspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArg 420  
Db 1293 AATGGAACCTCTGTGGGTAGTGTGAATCAACCATTTGGTGGCCTACACGTGATTGAG 1352  
Qy 421 GlyIleProValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrValTyr 440  
Db 1353 GGGCTACCCCTCTGGCTGACAGCAGCTGATGCATGGCCAGTGTGGCTGCCTACACCTAC 1412  
Qy 441 LysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLysLysIleArgVal 460  
Db 1413 CACCAGCATCTCTGGTTTCATCGGTACACGCGCGTAATTTGAAGAAGGTGCGGGTC 1472  
Qy 461 AspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGlnValValAspProGlyPro 480  
Db 1473 GATGGCTCTCAG--GATGCCAGCTATATGAACGGTCTCTGTGTGTCAGGGCAGCCCC 1529  
Qy 481 ValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGluArg 500  
Db 1530 ATACTTCGAGACCTGCTCTTCAGCCCTGACCCAGCACATCTACCTCTGAGTGAAG 1589  
Qy 501 GlnLeuThrArgValProValGluSerCysGlyGlnTyrGlnSerCysGlyGluCysLeu 520  
Db 1590 CAGGTGAGCCAACTCCCGGTGAGACCTGTGAGCAGTATCTGAGCTGCTGCATGCCCTG 1649  
Qy 521 GlySerGlyAspProHisCysGlyTrpCysValLeuHisAsnThrCysThrArgLysGlu 540  
Db 1650 GGCTCAGGGGACCCACACTGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1709  
Qy 541 ArgCysGluArgSerLysGluProArgArgPheAlaSerGluMetLysGlnCysValArg 560  
Db 1710 GCCTGTCCAGGTGCTCTGCCCCACATGGCTTTGCAGAGAGCTGAGCAAAATGTATACAG 1769  
Qy 561 LeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuLeuValLeuGlu 580  
Db 1770 GTGCGGGTCCGTCCTCCCAATATGTCTCAGTGACATCTCTCTGGGTGTCAGCTGACCC 1829  
Qy 581 ThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSerGlu 600  
Db 1830 ATGCCCAAGTCCGACACCTGAGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1889  
Qy 601 MetAspGlyLeuValVal--GlyAsnGlnIleGlnCysTyrSerProAlaAlaLysGlu 619  
Db 1890 AGTGAGGCTATTCTGTGCCCCCTGGAGAGTGTACGATGCTTTCACCATCCCTTCAGGAG 1949  
Qy 620 ValProArgIleIleThrGluAsnGlyAspHisValValGlnLeuGlnLeuLysSer 639  
Db 1950 CTCACACACTTACAGAGGGCATGGGCCACTACTACTGTGGGCTTCAGCTGCTCTCC 2009  
Qy 640 LysGluThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerValHis 659  
Db 2010 ATGGAGACTGGTGTGAGGTTGCTGGGGTTGACTTGTCTTCTACAACTGCAGTGCCTC 2069  
Qy 660 AsnSerCysLeuSerCysValGluSerProTyrArgCysHisTrpCysLysTyrArgHis 679

Db 2070 CAGTCGTGTATGTCTCTGTGTGGCAGCCCTTACCCCTGCCACTGTGTAAAGTACCGTCA 2129  
Qy 680 ValCysThrHisAspProLysThrCysSerPheGlnGluGlyArgValLysLeuProGlu 699  
Db 2130 GTGTGTACCACCCACCCACAGTGTCTTTCCAGGAGGCGAGGTCCACAGCCCTGAG 2189  
Qy 700 AspCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysPro 719  
Db 2190 GGCTGCCCTGAGATCTCTCAAGGGGACCTTCTGATTCTCTGTGGTGTCTATGACGCT 2249  
Qy 720 IleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCys 739  
Db 2250 CTAACCTCGGGGCTAAGAACCTCCACAGCTCAGTCCGACACAGAAATATTAATGCT 2309  
Qy 740 IleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArgPheAsnSerSer 759  
Db 2310 GTGGTCCGTGTACAGGAGCGCAACATCGGGTACCTGTCAGTGGCTTTAAACAGCAGC 2369  
Qy 760 ValGlnCysGlnAsnThrSerTyrGluGlyMetGluIleAsnAsnLeuProVal 779  
Db 2370 GTGCAGTCCCAAGATCTCTCTCTATGAAGGTGATGAGTTTGGTGACACCGCACTG 2429  
Qy 780 GluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsnProAlaGlnAsnLysVal 799  
Db 2430 GACTCTCTGTGGTGGGATTTCCCAATGATAAGCCCTCCTAGTTCGAGCC 2489  
Qy 800 HisLeuTyrLysCysGlyAlaMetArgGluSerCysGlyLeuCysLeuLysAlaAspPro 819  
Db 2490 CTTCTTTACAAGTCTGGGCTCAGGGCTAGTGTGGCTCTGCTCAAGGCTCATCC 2549  
Qy 820 AspPheAlaCysGlyTrpCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysPro 839  
Db 2550 CGATTCAACTGTGGTGTGTGTCTCAGACACAGGTGCCAGTCTCAGGGCTCACTGCCCA 2609  
Qy 840 AlaGlnGluSerGlnTrpLeuGluLeuSerGlyAlaLysSerLysCysThrAsnProArg 859  
Db 2610 GCTCCCAAAAGTAAGTGGATGACCCCAAGGAGGTGGCCGATGAGCCATCCCCGA 2669  
Qy 860 IleThrGlnIleLeuProValThrGlyProArgGluGlyGlyThrLysValThrIleArg 879  
Db 2670 ATCACCAGATTTCATCCACTTACAGGCCCAAGGAGGTGGCAGCCGGTCAACATTGTG 2729  
Qy 880 GlyGluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyVal 899  
Db 2730 GGTGAGAACCTGGGCTCACTTCCCGTGGTGGC-----CTTCGAGTAGTGGTGA 2783  
Qy 900 GluCysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMetGly 919  
Db 2784 CGTTGCAACTCCATTCCTCCACCGAATATGTCTAGTGTGAAGGATCGTATGTGAGATGGAG 2843  
Qy 920 GluAla---LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArg 938  
Db 2844 GAATCGTGGTGGCCAGCCACCCTGGGCTGTGTGAGCTGTGTAGGGGATTGTCT 2903  
Qy 939 ProGluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAsp 958  
Db 2904 GCTGACTTCCGCACACAGTCCAGCACTCTACAGCTTTGTGACCCCTACATTTGACCGT 2963  
Qy 959 LeuLysProSerArgGlyProMetSerGlyThrGlnValThrIleThrGlyThrAsn 978  
Db 2964 GTGAGTCCCTCTGGGGCCAGCTCTGGAGGCACTCGGCTTACCATCTCTGGAATTTCT 3023  
Qy 979 LeuAsnAlaGlySerAsnValValMetPheGlyLysGlnProCysLeuPheHisArg 998  
Db 3024 CTGGATCCCGCAGCAGGGTCAAGTGTATTAAGAGATGGCGAGTGGCAGTTGTGAGG 3083  
Qy 999 ArgSerProSerTyrIleValCys---AsnThrThrSerSerAspGluValLeuGluMet 1017  
Db 3034 AGAGATCCGAGGCAATCGTGTATCTCAGCTGTCTCCACCTGGGCCCCAGTCAGAGC 3143  
Qy 1018 LysValSerValGlnValAspArgAlaLysIle---HisGlnAspLeuValPheGlnTyr 1036  
Db 3144 CCTATCACCTTGGCATCGATCCCAACACTCTCCAACTGGGAGTCTATCTATACCTAC 3203



```

Db      5346  TTTGTTTTCACATCATAGATACATACAGACCGCTGTTGTGACGTGGTGGCCAG 5405
QY      1777  ThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsp 1796
Db      5406  ACCTTCATGAGTCCTGTTCTACATCAGACACCGCTGGCAGGAGCTACCTTCCTCAAC 5465
QY      1797  LysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSer 1816
Db      5466  AAGCTCCTCTATGCCAAGGATATTCCTCAATACCAAGAGCTGGTGGAGAGTACTATCGA 5525
QY      1817  AspIleGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluGln 1836
Db      5526  GATATCGCAAGATGCGATCCATCATGACCTGACGAGATGACCGCTACTTGTAGTGAGCAG 5585
QY      1837  SerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyr 1856
Db      5586  TCCCGCTCCATGCTAATGACTTCATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAAT 5645
QY      1857  ValGlyLysTyrSerGluGluLeuLeuGlyProLeuAspHisAspGlnCysGlyLys 1876
Db      5646  GTCACCAAGTACCGTCAGGAGATCCTCACCTCGCTGACCGAGATGCTTGTGCGGAAG 5705
QY      1877  GlnLysLeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerLeuAspSer 1894
Db      5706  CACAAGCTTCGAGAGAAGCGCTGAGAGATCATCACCTGCTGCTGCTGCTGCTGCTGCTGCT 5759

```

## RESULT 9

```

HSNOVPROT      5263 bp      mRNA      linear      PRI 01-OCT-1999
LOCUS      Homo sapiens mRNA for partial NOV/plexin-A1 protein.
DEFINITION      X87832
ACCESSION      X87832
VERSION      NOV/plexin-A1 gene: transmembrane receptor.
KEYWORDS      Homo sapiens.
SOURCE      Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 5263)
AUTHORS      Maestrini, E., Tamagnone, L., Longati, P., Cremona, O., Gullisano, M.,
Bione, S., Tamagnini, F., Neel, B.G., Toniolo, D. and Comoglio, P.M.
TITLE      A family of transmembrane proteins with homology to the
MET-hepatocyte growth factor receptor
Proc. Natl. Acad. Sci. U.S.A. 93 (2), 674-678 (1996)
JOURNAL      96149362
MEDLINE      8570614
PUBMED
REFERENCE      2 (bases 1 to 5263)
AUTHORS      Tamagnone, L., Artigiani, S., Chen, H., He, Z., Ming, G.L., Song, H.,
Chedotal, A., Winberg, M.L., Goodman, C.S., Poo, M., Tessier-Lavigne, M.
and Comoglio, P.M.
TITLE      Plexins are a large family of receptors for transmembrane,
secreted, and GPI-anchored semaphorins in vertebrates
Cell 99 (1), 71-80 (1999)
JOURNAL      99449305
MEDLINE      10520995
PUBMED
REFERENCE      3 (bases 1 to 5263)
AUTHORS      Toniolo, D.
TITLE      Direct Submission
Submitted (08-JUN-1995) D. Toniolo, Istituto di Genetica
Biochimica ed Evoluzionistica, CNR, Via Abbiategrasso 207, 27100
Pavia, ITALY
REMARK      Revised by [4]
REFERENCE      4 (bases 1 to 5263)
AUTHORS      Tamagnone, L.
TITLE      Direct Submission
JOURNAL      Submitted (27-JUN-1998) L. Tamagnone, University of Torino, IRCC,
SP 142, 10060 Candolo, Torino, ITALY
COMMENT      On Oct 4, 1999 this sequence version replaced gi:1247575.
FEATURES      Location/Qualifiers
1..5263
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
source

```

```

/map="q21-qter"
/tissue_type="skeletal muscle"
/clone_lib="lambda GT10"
1..5263
/gene="NOV/plexin-A1"
<1..5263
/gene="NOV/plexin-A1"
/codon_start=2
/product="NOV/plexin-A1 protein"
/protein_id="CAB57274.1"
/db_xref="GI:6010217"
/translation="GMWAEAGLPRAGGSGQPPFRFTFGSDWGLTLLHLLVHEQGTGVVVG
ANRYIKSGNLTLLRAHVTGPEVDNEKCYPPPVQSQCPHGLGNTDNNKLLLLDAA
NRLLAGGSASQICQSLRLDLXFLKLEPHRKEHYLSSVOEAGIAGVLPAGPGGQ
AKFVGTIDGKSEYFPTLSSLMANEDADMFGVYODEFVSQKLPSTLSRPF
AFDIYVYFSEQFVYLLTLOLDTQSPDAGHEFTSKIVRLCDDPKFYVVEF
PIGCEAGVYRLVQDALYSRPRALAHOLGLAEDEDVLTFTVPAQOKRNPRESA
LCFLTRAIKERIKERIOSCYRGEGKLSLPWLNLKELGINSPLQIDDDFGQDFNPQ
LGGVTITEGTLFVDKDLTAAYDYRGRTVVFAGTRSGRIKILYDLISNPGRPA
LAYESVAAQEGSPILRLDLVSPNHYLYAMTEQVTRVPSVCQVTSCELCGSRDP
HCGWCLVHSGISRRDADERADPQFAADLCQVLTQVPRNVSVTMSOVPLVLOAWN
VPDLSAGVSCSFEDTSESVLEDRGRIHCRSPSAREVAPITRGQDORVVKLYXKSE
TGKFAVDVFYFVYNSVHQSCLSVNGSPFCHWCKYRHVCTHNVDACAELEGRNVSE
DCQILPSTQIYVPGVVKPITLAARNLPQPGQSGRGYECLEPHIIPGSPARVATLRNS
SSLCQNSSYSYEGNDVSDLPVNLVSVNGNFVIDNPQIAHLKPCALRSCGLCL
KADPRFCGWCVAERRCSLRHCAADTPASMHARHGSSRCQIVCEIGDASSVRAHADL
GTRLTITCENGLRPEDVRLGVGVKVLCSVSEVISAQIVCEIGDASSVRAHADL
VEVCVRDCSPHYRALSPKRFVTFTFVRVSPSRGPGTGWIGEGSHLNAGSDVAV
SYGGRPCSFNRNSREICLTPPGQSPGSAPIIININRAQLNPVKNYKTEDPTILR
IDPWSINSGGTLLTVTGNLATVREPRIKAYGGIERENGCLVYNDTTCRAPSA
NPVRSPPELGERPDELGFVNDVRLSLNLNSTFLYVDPVLEPLSTGLLELKPSP
LILKGRNLLPPAGPNSRNYTLVIGTPTCTVTSYQTLCEAPNTGQHKVTVRAGGF
EFGPTLQVSDLTLPALVIGGGGLLVIVAVLTAYRKRSADATLRLQLQ
MDNLSRYALECKEAFALQTDIHELTNDLDCAGTFLDYTYAMRVLPGLDEHPVL
KEMEVAQNVKSLTFLGQLLKHFLLFIPTLEAQRSEMRDGNVSLIMTALQGE
MEYATGVULKQLSDLEKNLKHLLRRSTSEVAEKMLTNWFTFLYKFLKECAG
EPFLMYCAIQKQEKGPIDAITGEARYSLSDKLTIXQIDYKTLTLCNVNENAP
EVPYKGLDQVTVQAKELLDAAVGVPSYQRPKADMDLEWRQGRMARIIQDEEDVT
TKIDNDKRLNTLAHYQVTDGSSVALVPKQTSAYNTSNSTFKLSRYSLEMTASS
BLSRSTPMITPDLESCTKLWHLVKNHDLQREGDRGSKMVSEIYLRLLATKCTL
QKQVDDLETFETFAHRSALPLAIKYMFDLDEQADKHQIHADVVRHTWKSCLPLR
FWNVVKNPQFVDLHKNSITDCLSV"
BASE COUNT      1034 a 1723 c 1501 g 939 t 6 others
ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      5263
Score:          5847.50      Matches:      1114
Percent Similarity: 77.29%      Conservative: 237
Best Local Similarity: 63.73%      Mismatches: 378
Query Match:      58.53%      Indels:      19
DB:              9      Gaps:      12
US-09-964-956-13 (1-1896) x HSNOVPROT (1-5263)
QY      37  GlnArgSerPheValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValVal 56
Db      47  CAGCCCCCTTCCGACCTTCTCGGCGACGAC---TGGGGCCTCACCCACCTATTGGTG 103
QY      57  AspGluArgThrGlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSer 76
Db      104  CATGACGACAGACGCGAGGTGTATGTGGCGCAGTGAACCGCATCTATAGCTGTCGGGG 163
QY      77  AspleuLysValLeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyr 96
Db      164  AACCTGACACTGCTCGGGCCCGACGTCACGGGCCCTGTGGAGGACAAACGAGAGTCTAC 223
QY      97  ProProArgIleValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLys 116
Db      224  CCGCGGCCCANCGTCAGTCTCTGCCCCCGCCGCGCTGGGCAATACTGCACACGCTCAACAAG 283
QY      117  MetLeuLeuIleAspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGly 136

```

Qy	490	AspHisGluGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSer	509
Db	1418	AACCACAGCATTACCTNTACGCCATGACTGAAAGCAGGTGACGGCGGTGCCTGTGGAGAGC	1477
Qy	510	CysGlyGlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTyr	529
Db	1478	TGTGTGCAGTACAGGTCTCTGTGAGCTGTGTCTGGGGTTCACGGGACCCCACTGTGGCTGG	1537
Qy	530	CysValLeuHisAsnThrCysThrArgLysGluArgCysGluArgSerLysGluProArg	549
Db	1538	TGTGTCTCTGCACAGCATCTGCTCGCGCGGGACGCTGTGAGCGAGCAGACGAGCCCGCAG	1597
Qy	550	ArgPheAlaSerGluMetLysGlnCysValArgLeuThrValHisProAsnAsnIleSer	569
Db	1598	CGCTTGTCTCGGACCTGCTGCAGTGTGTGACGTGACTGTGCAGCCCGCCCAATGTGTCT	1657
Qy	570	ValSerGlnTyrAsnValLeuValLeuGluThrTyrAsnValProGluLeuSerAla	589
Db	1658	GTCCACATGTCGCCAGGTCCCACCTGTGCTGCAGGCCCTGGAACGTGCCTGACCTCTCAGCT	1717
Qy	590	GlyValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValGlyAsnGln	609
Db	1718	GGCGTCAACTGCTCTCTTCGAGGACTTCACGGAACTCTGAGAGCGTCTCTGGAGATGGCCGG	1777
Qy	610	IleGlnCysTyrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAsp	629
Db	1778	ATCCACTGCGCTCACCTCTCGCCCGGAGGTGGCGCCCATCAGCGGGGCCAGGGAGAC	1837
Qy	630	HisHisValValGlnLeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerThr	649
Db	1838	CAGCGGTGTGTAACACTCTACNTAAAGTCCAAGGAGACAGGGAAGAAGTTTGCCTCTGTG	1897
Qy	650	SerPheValPheTyrAsnCysSerValHisAsnSerCysLeuSerCysValGluSerPro	669
Db	1898	GACTTCGTCTTCTACAACTCGAGCGTCCACAGTCTCGCTGCTCTGTGTCAACGGCTCC	1957
Qy	670	TyrArgCysHisTrpCysLysTyrArgHisValCysThrHisAspProLysThrCysSer	689
Db	1958	TTTCCCTGCCACTGGTCAAAATACCGCACCGTGTGCACACACACACACACACACACAC	2017
Qy	690	PheGlnGluGlyArgValLysLeuProGluAspCysProGlnLeuLeuArgValAspLys	709
Db	2018	TTCTCTGAGGGCGGTGCAACGTGTCTGAGGACTGCCACAGATCTCTGCCCTCCACGCAG	2077
Qy	710	IleLeuValProValGluValIleLysProIleThrLeuLysAlaLysAsnLeuProGln	729
Db	2078	ATCTACGTGCCAGTGGGAGTGGTAAACCCATCACCTGGCGGCACACGAACTGCCACAG	2137
Qy	730	ProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGluGlnArg	749
Db	2138	CCACAGTCAAGCCACGCTGGATATGATGGCTCTTTCACACTCCCGGCAGCCCGCCCGCT	2197
Qy	750	ValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThrSerTyrSerTyr	769
Db	2198	GTCCACGCCCTGCGCTTCAACAGTCCAGCTCGAGTGCAGGATTCCTCGTACTCTCTAC	2257
Qy	770	GluGlyMetGluIleAsnAsnLeuProValGluLeuLeuThrValValThrAsnGlyHisPhe	789
Db	2258	GAGGGGAACATGTACGGACCTGCCAGTGAACCTGTCTGAGTCGTGTGGACGGCAACTTT	2317
Qy	790	AsnIleAspAsnProAlaGlnAsnLysValHisLeuTyrLysCysGlyAlaMetArgGlu	809
Db	2318	GTCAATGACAAACCCACAGACATCCAGGCGGACCTCTACAGTGCCTCGGCGCTCGGCGAG	2377
Qy	810	SerCysGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTyrCysGlnGlyPro	829
Db	2378	AGCTGCGGCTCTGCCTCAAGGCGGACCCGCGCTTCGAGTGCAGATGCTGCTGCGCCGAG	2437
Qy	830	GlyGlnCysThrLeuArgGlnHisCysProAlaGln---GluSerGlnTrpLeuGluLeu	848
Db	2438	CGCCGNTGCTCCCTGCGACACCACTGCGCTGCTGCCGACACACCTGCATGCTGGATGCACGCG	2497

QY 849 SerGlyAlaLysSerLysCysThrAsnProArgIleThrLeuIlelleProValThrGly 868  
DB 2498 CGTCACGGCAGCAGTGCCTGCACCGACCCCAAGACTCTCAAGCTGCCCGGAGCGGC 2557  
QY 869 ProArgGluGlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuPheArg 888  
DB 2558 CCGAGGCGAGGGCGGACCGGGCTCACTATACAGAGCGAGAACCTGGCGCTCGCAATCGAA 2617  
QY 889 AspIleAlaSerHisValLysValAlaGlyValGluCysSerProLeuValAspGlyTyr 908  
DB 2618 GAGGTGCTGGCGCTGGCGGTGGCAGAGTGTGTGCAGCCCTGTGGAGACGAGTAC 2677  
QY 909 IleProAlaGluGlnIleValCysGluMetGlyGluAlaLysProSerGln---HisAla 927  
DB 2678 ATCAGTGCAGGACATGCTGTGTGAGATCGGGGACGCCAGCTCCGTGGTGGCCCATGAC 2737  
QY 928 GlyPheValGluIleCysValAlaValCysArgProGluPheMetAlaArgSerSerGln 947  
DB 2738 GCCCTGGTGGAGGTGTGTGGGGACTGCTCACACACTACCGCGCCCTGTCCACCTAAG 2797  
QY 948 LeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMetSer 967  
DB 2798 CGCTTCACCTTCGTGCACACCACTTCTACCGTGTGAGCCCTCCCGTGGGCCCTCTGTCA 2857  
QY 968 GlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnValVal 987  
DB 2858 GGGGCGACCTGGATTGCGATCGAGGGAAGCACCTCAACGAGCAGGAGTGTGTGTGTG 2917  
QY 988 MetPheGlyGlnProCysLeuPheHisArgArgSerProSerTyrIleValCysAsn 1007  
DB 2918 TCGGTGGTGGCGCGCGCTGCTTCTCTGGAGGAATCCCGTGGAGATCCCGGTGCCCTG 2977  
QY 1008 ThrThrSerSerAspGluValLeuGluMetLysValSerValGlnValAspArgAlaLys 1027  
DB 2978 ACACCCCCGGCAGAGCCCTGCGCGCTGCCATCATCATCAACATCAACCGCGCCAG 3037  
QY 1028 Ile---HisGlnAspLeuValPheGlnTyrValGluAspProThrIleValArgIleGlu 1046  
DB 3038 CTCACCAACCTCGAGTGAAGTAACTACACCGAGGAGCCCACTCATCTCTGAGGATCGAC 3097  
QY 1047 ProGluTrpSerIleValSerGlyAsnThrProIleAlaValTrpGlyThrHisLeuAsp 1066  
DB 3098 CCGAGTGGAGCATCAACAGCGGTGGGACCTCTCGTACGCTACAGCCACCACTTGCC 3157  
QY 1067 LeuIleGlnAsnProGlnIleArgAlaLysHisGlyGlyLysGluHisIleAsnIleCys 1086  
DB 3158 ACTGTCCGTGAACCCGAATCCGGGCCAAGATATGAGGACATTGAGAGGAGACGGCTGC 3217  
QY 1087 GluValLeuAsnAlaThrGluMetThrCysGlnAlaProAlaLeuAlaLeuGlyProAsp 1106  
DB 3218 CTGGTGTACAATGACACCACTGATGGTATGCGCGCGCCGCTGTGGCCAAACCTGTGCGC 3277  
QY 1107 HisGlnSerAspLeuThrGluArgProGluGluPheGlyPheIleLeuAspAsnValGln 1126  
DB 3278 AGCCACCAAGAGCTGGGGAGCGCGGATGAGCTGGGTTCGTATGGACACAGTGGC 3337  
QY 1127 SerLeuLeuIleLeuAsnLysThrAsnPheThrTyrTyrProAsnProValPheGluAla 1146  
DB 3338 TCCCTGCTGTGCTCAACTCCACTCTCTCTACTACTACCTGACCCCGTACTGGACCA 3397  
QY 1147 PheGlyProSerGlyIleLeuGluLeuLysProGlyThrProIleLeuLysGlyLys 1166  
DB 3398 CTCAGCCCCACTGGCTGTGGAGCTGAAGCCAGCCAGCTCCCACTCATCTCAAGGGCCGG 3457  
QY 1167 AsnLeuIleProProValAlaGlyGlyAsnValLysLeuAsnTyrThrValLeuValGly 1186  
DB 3458 AACCTCTGCCACCT---GCACCCGCAACTCCCGACTCAACTACACGGTGTCTCATGGC 3514  
QY 1187 GluLysProCysThrValThrValSerAspValGlnLeuLeuCysGluSerProAsnLeu 1206  
DB 3515 TCCACACCTGTACCTTACCGTGTGCGAGACCACTGCTGTGCGAGGGCCCAACCTC 3574  
QY 1207 IleGlyArgHisLysValMetAlaArgValGlyGlyMetGluTyrSerProGlyMetVal 1226

DB 3575 ACTGGCAGCAACAAGTCCACGGTGGCGCAGGTGCTTTCGAGTCTTCGCGAGGACACATG 3634  
QY 1227 TyrIleAlaProAspSerProLeuSerLeuProAlaIleValSerIleAlaValAlaGly 1246  
DB 3635 CAGGTGTACTCGGACAGCCTGCTACGCTGCTGCCATTTGTGGGATTTGGCGGAGCGGC 3694  
QY 1247 GlyLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrLysArgLysSerArgGlu 1266  
DB 3695 GGTCTCTGCTGCTGGTGCATCGTGCCTGCTAATCGCTACAAGCGCAAGTCACGAGAT 3754  
QY 1267 SerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAla 1286  
DB 3755 GCTCAGCGCACACTCAAGCGCTGCAGCTCCAGATGGACAACCTGGAGTCCCGCGTGGCC 3814  
QY 1287 LeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIleHisGluLeuThrSerAsp 1306  
DB 3814 CTCAATGCAAGGAAGCCTTTGSCAGAGCTGCAGACAGATCCACGAGTGCACCAATGAC 3874  
QY 1307 LeuAspGlyAlaGlyIleProPheLeuAspTyrArgThrTyrThrMetArgValLeuPhe 1326  
DB 3875 CTGGACGGTGGCGCATCCCTTCTTCACTACCGACATATGCCATGCGGGTGTCTTT 3934  
QY 1327 ProGlyIleGluAspHisProValLeuArgAspLeuGluValProGlyTyrArgGlnGlu 1346  
DB 3935 CCTGGATCGAGGACCACTTGTCTCAAGGAGATGGAGTG-----CAGGCC 3982  
QY 1347 ArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIleAsnAsnLysValPheLeuLeu 1366  
DB 3983 AATGTGGAAGATCGCTGCACACTGTTGGGAGCTGTGCACCAAGAACACTTCCTGCTG 4042  
QY 1367 SerPheIleArgThrLeuGluSerGlnArgSerPheSerMetArgAspArgGlyAsnVal 1386  
DB 4043 ACCTTCATCCGACGCTGCGAGGACGACGAGCTTCTCCATCGCGGACCGCGGGAATGTG 4102  
QY 1387 AlaSerLeuIleMetThrValLeuGlnSerLysLeuGluTyrAlaThrAspValLeuLys 1406  
DB 4103 CCTCGCTCATCATGACGCGCTGCGAGGCGAGATGGAATACGACAGCGCTGTCTCAAG 4162  
QY 1407 GlnLeuLeuAlaAspLeuIleAspLysAsnLeuGluSerLysAsnHisProLysLeuLeu 1426  
DB 4163 CAGTGTCTTCGACCTCATCGAAGAACCTGGAGAGCAAGAACCCCAAGCTGTCTCAAG 4222  
QY 1427 LeuArgArgThrGluSerValAlaGluLysMetLeuThrAsnTrpPheThrPheLeuLeu 1446  
DB 4223 CTGCGCGGACTGAGTGTGCGCAGAGAAGATGTAACTTAACCTTCACTTCTCTTCTG 4282  
QY 1447 TyrLysPheLeuLysGluCysAlaGlyGluProLeuPheSerLeuPheCysAlaIleLys 1466  
DB 4283 TATAAGTTCTCAAGGAGTGGCTGGGAGCGCTGTTTCATGCTGTACTGTGCGCCATCAAG 4342  
QY 1467 GlnGlnMetGluLysGlyProIleAspAlaIleThrGlyGluAlaArgTyrSerLeuSer 1486  
DB 4343 CAGCAGATGGAGAGGGGCCCATTCACGCCATACGCGGTGAGGACGCTACTCCCTGAGT 4402  
QY 1487 GluAspLysLeuIleArgGlnGlnIleAspTyrLysThrLeuValLeuSerCysValSer 1506  
DB 4403 GAGGACAAAGCTCATCCGCGCNCAGATTGACTACAAGACACTGACCTGAACTGTGTGAAC 4462  
QY 1507 ProAspAsnAlaAsnSerProGluValProValLysIleLeuAsnCysAspThrIleThr 1526  
DB 4463 CCTGAGAATGAGAATGCACTCAGGTGCGGTGAAGGGGCTGGACTGTGACACGCTCACC 4522  
QY 1527 GlnValLysGluLysIleLeuAspAlaIlePheLysAsnValProCysSerHisArgPro 1546  
DB 4523 CAGGCCAAGGAGAGCTGCTGACCTGCTCAAGGGCGTCCCTTACTCCCGAGGGGCC 4582  
QY 1547 LysAlaAlaAspMetAspLeuGluTrpArgGlnGlySerGlyAlaArgMetIleLeuGln 1566  
DB 4583 AAGCGCGGACATGACCTGGAGTGGCGCCAGGCGCGCATGCGCGCATCATCTCTGAC 4642  
QY 1567 AspGluAspIleThrThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHis 1586



Db 62 GGGGTCACGAGTACGAGTGTCTGTGCTCAAGGACGGAGCCCAATCCTCCGGACATG 121  
QY 486 AlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGluArgGlnLeuThrArgVal 505  
Db 122 GCCTTCCATTGATCAGCGCTACCTGTAGGTCTATGCTGAGAGACAGGTCACCAAGGTC 181  
QY 506 ProValGluSerCysGlyGlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspPro 525  
Db 182 CCGGTGAGTCAATGTAGCAGTATACGACTGTGGGAGTGCCTGAGCTCTGGGGACCT 241  
QY 526 HisCysGlyTrpCysValLeuHisAsnThrCysThrArgLysGluArgCysGluArgSer 545  
Db 242 CACTGTGGCTGGTGGCTGTCACACATGTGCTCCGCGAGGAGCAAAATGCCAACAGGCC 301  
QY 546 LysGluProArgArgPheAlaSerGluMetLysGlnCysValArgLeuThrValHisPro 565  
Db 302 TGGGAACCTAATCGATTGCTGCCAGCATCAGCCAGTGTGTGAGCCTTGGCAGTGATCCC 361  
QY 566 AsnAsnIleSerValSerGlnTyrAsnValLeuLeuValLeuGluThrTyrAsnValPro 585  
Db 362 AGCAGCATCTCAGTATCTGAGCACAGCCGGTGTGCTTAGCCTGTGAGTGTGATGCTCCT 421  
QY 586 GluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuVal 605  
Db 422 GATCTATCTCGGGGTATCGGCTGTGCTTTGGGAACCCGACAGAGTGGAGGGCAGGTG 481  
QY 606 ValGlyAsnGlnIleGlnCysTyrSerProAlaAlaLysGluValProArgIleThr 625  
Db 482 TCCGGGACCGCAGTCTGCTGCTACCTGCGGCCCAAGGATGTCCT---GTCTATCCG 538  
QY 626 GluAsnGlyAspHisValValGlnLeuGlnLeuLysSerLysGluThrGlyMetThr 645  
Db 539 CTGGATCAAGACTGGTGTGGCTGGAGCTACAGCTGAGGTCCCAAGGACAGAGGAAGATA 598  
QY 646 PheAlaSerThrSerPheValPheTyrAsnCysSerValHisAsnSerCysLeuSerCys 665  
Db 599 TTTGTCAGCACCCAGGTTCAAGTTTACACTGAGTGGCCCAACACCGTGCCTGTCTGT 658  
QY 666 ValGluSerProTyrArgCysHisTrpCysLysTyrArgHisValCysThrHisAspPro 685  
Db 659 GTCAACAGCGCCTTCCGCTGCCATTGGTGCAGTACCGCAACCTCTGCACTCATGACCCC 718  
QY 686 LysThrCysSerPheGlnGluGlyArgValLysLeuProGluAspCysProGlnLeuLeu 705  
Db 719 ACCACTGCTCTCTCCAGGAGGCGGATCAATATTTTCAGAGGACTCTCCCGAGCTGGT 778  
QY 706 ArgValAspLysIleLeuValProValGluValIleLysProIleThrLeuLysAlaLys 725  
Db 779 CCCACAGAGGAGATCTTGATTCAGTGGGAGGTAAAGCAATCACTTAAGGCCGGA 838  
QY 726 AsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGly 745  
Db 839 AATCTGCCCGCAGCGCAGTCCGCGCAGGAGGCTATGAGTGTGCTCTCAACATACAGGA 898  
QY 746 SerGluGlnArgValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThr 765  
Db 899 GCATCCACCGCGGTCCCGCTCTCGCGCTTCAACAGCTCCAGCCTTCAAGTTCAGCAACAG 958  
QY 766 SerTyrSerTyrGluGlyMetGluIleAsnAsnLeuProValGluLeuThrValValTrp 785  
Db 959 TCGTACAGTATGATGGCATGGACATCAGCAATCTGCGCGTGGATTCGCTGGTGTGG 1018  
QY 786 AsnGlyHisPheAsnIleAspAsnProAlaGlnAsnLysValHisLeuTyrLysCysGly 805  
Db 1019 AACGGCAATTTTCATATTGACAACCTCAGGACCTGAAAGTCCATCTCTACAAAGTGCA 1078  
QY 806 AlaMetArgGluSerCysGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTrp 825  
Db 1079 GCCACGCGGAGAGCTGCGGCTCTGCTTCAAGCGCAGCCGGAAGTTTGGTGTGGCTGG 1138  
QY 826 CysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysProAlaGlnLeuSerGlnTrp 845  
Db 1139 TGCAGCGGCGAGCGGAGGTGCACCCCTCCACCACTGTATACCAGCCCTTCCAGCCCTGG 1198

QY 846 LeuGluLeuSerGlyAlaLysSerLysCysThrAsnProArgIleThrGluIleIlePro 865  
Db 1199 CTCGACTGGTCCAGCCCAANTGTCAAGTGTCCAACTCCAACTCAACAGAGATTTTGACG 1258  
QY 866 ValThrGlyProArgGluGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeu 895  
Db 1259 GTGTCTGGACGGCCGAGGAGGAGGAGTGTGATTTGGCGAGTGTAGCCAGAGTTCATGACGAAGTCC 1318  
QY 886 GluPheArgAspIleAlaSerHisValLysValAlaGlyValIleCysSerProLeuVal 905  
Db 1319 GACTTCTCCGAGATCCGCCACCATGTGAGGTGGTGGGTCCTGTCAGCCGCCCTCCCA 1378  
QY 906 AspGlyTyrIleProAlaGluGlnIleValCysGluMetGlyGluAlaLysProSerGln 925  
Db 1379 GGGGAATACATCATCGCTGAGCAGATGTCTGTGAGATGGCCATCCCTCGTGGGAACC 1438  
QY 926 HisAlaGlyPheValGluIleCysValAlaValCysArgProGluPheMetAlaArgSer 945  
Db 1439 ACCTCCGGCGCAGTACGGCTGTGTATTTGGCGAGTGTAGCCAGAGTTCATGACGAAGTCC 1498  
QY 946 SerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyPro 965  
Db 1499 CATCAGCAGTACACCTCTGCGAACCTTCTGTGCTCACTCAACCCATCCGAGGTCCC 1558  
QY 966 MetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnVal 985  
Db 1559 GAGTCAGGAGGACACTATGTGACCATACCGGCCATTACCTTGGGCTGGGACGCGTG 1618  
QY 986 ValValMetPheGlyGlyGlnProCysLeuPheHisArgArgSerProSerTyrIleVal 1005  
Db 1619 GCAGTCTACCTGGGCAACGACGAGCTGCGAGTTCTACGGGAGGTCAATGATGAGATCGTG 1678  
QY 1006 Cys---AsnThrThrSerSerAspGluLeuLeuMetLysValSerValGlnValAsp 1024  
Db 1679 TGTGCTCACCCCAATCACTCAATGGCTGGCCGGTCCCTGTTCTGTGAGTGTGCGAC 1738  
QY 1025 ArgAlaLysIleHisGlnAspLeuValPheGlnTyrValIleAspProThrIleValArg 1044  
Db 1739 CGAGCCCATGTGATAGCAACCTGAGTGTGAGTACATAGATGACCTCGGGTCCAGCGC 1798  
QY 1045 IleGluProGluTrpSerIleValSerGlyAsnThrProIleAlaValTrpGlyThrHis 1064  
Db 1799 ATCGAGCCAGAGTGGAGCATTCGCCAGTGGCCACACACCCCTGACCATCAGGTTCAAC 1858  
QY 1065 LeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGlyGlyLysGluHisIleAsn 1084  
Db 1859 CTGGATGTCATTCAGGAGCCAAAGGATCCGAGTCAAAATTCATGGCAAGAAATCTGCAAT 1918  
QY 1085 IleCysGluValLeuAsnAlaThrGluMetThrCysGlnAlaProAlaLeuGly 1104  
Db 1919 GTGTGTAAGTTGTCAACACACACCCCTCACCTGCTGGCACCTCTCTGACCCAGGAC 1978  
QY 1105 ProAspHisGlnSerAspLeuThrGluArgProGluGluPheGlyPheIleLeuAspAsn 1124  
Db 1979 TACCGCCCTGGCCCTGGACACTGTGAAACGCCAGATGAGTTTGGATTTCTTTAAACAAT 2038  
QY 1125 ValGlnSerLeuLeuIleLeuAsnLysThrAsnPheThrTyrProAsnProValPhe 1144  
Db 2039 GTCCAACTCTTGTAAATTTACACACACCAAGTTTACTACTACCCCAACCCGACCTTT 2098  
QY 1145 GluAlaPheGlyProSerGlyIleLeuGluLeuLysProGlyThrProIleIleLeuLys 1164  
Db 2099 GAACGTCTTAGCCCTACTGGAGTCTTGGATCAAAAGCCAGGATCGGCCCATCTCTTGAAG 2158  
QY 1165 GlyLysAsnLeuIleProProValAlaGlyGlyAsnValLysLeuAsnTyrThrValLeu 1184  
Db 2159 GGGAAAAACCTCTGCGCTCTGCGCTGGAGG---GCCAAACTCAACTACACTGTGCTC 2215  
QY 1185 ValGlyGluLysProCysThrValThrValSerAspValGlnLeuLysCysGluSerPro 1204  
Db 2216 ATCGGAGAGACCCCTTGTGCTCACCGTATCTGAGACCCAGCTTCTCTGCGAGCCCTGCC 2275



\* 1 64436: contig of 64436 bp in length  
 \* 64437 64536: gap of unknown length  
 \* 64537 198653: contig of 134117 bp in length.

## FEATURES

source Location/Qualifiers

1..198653  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="X"  
 /clone="RP21-430D6, RP21-367E1, RP21-561C20, RP21-643D22, RP21-17J5, RP21-412F21"  
 /clone="RP21-643D22"  
 /clone\_lib="RPCI mouse PAC library 21"  
 <1..3262

source

/organism="Mus musculus"  
 /db\_xref="taxon:10090"

source

<1..64918  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"

source

<1..150048  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"

source

28437..158588  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"

source

86082..162305  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"

source

/clone="RP21-412F21"

BASE COUNT

54705 a 43176 c 45578 g 55094 t 100 others

ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 198653  
 Score: 4421.00 Matches: 1115  
 Percent Similarity: 37.80% Conservative: 292  
 Best Local Similarity: 29.96% Mismatches: 431  
 Query Match: 44.25% Indels: 1896  
 DB: 2 Gaps: 43

US-09-964-956-13 (1-1986) x AC091474 (1-198653)

QY 8 TrpThrCysLeuLeuSerHisLeuLeu-----MetValGly-----Met 20  
 DB 33378 TGGTCATGCCCATGCTCTGC-CTTCCTCCCATTCCTATTCTTCCACCATAGAGGGTGCCTG 33436  
 QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
 DB 33437 GGTAGCAGC-----AGGCCATT 33454  
 QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
 DB 33455 CGTACCTTTGTGTAAACAGATACCCT-CTGACTCACCTGGCTGTGCACCGAGTGACT 33511  
 QY 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80  
 DB 33512 GGGGAGGTGTTTGTAGTCCGCGTGAATCGAGTTTCAAGTGGCTCCAAACCTAACTGAG 33571  
 QY 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100  
 DB 33572 CTACGGGGCCCATCTCAGAGGGCCCATAGGACAATGCTCGCTGACCCACCTCCCTAGC 33631  
 QY 101 ValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeu 120  
 DB 33632 ATCCGTGTGTGCTCCACCGCCCTGGTCCCTGTGGACAATGTGAACAGCTGCTCCTCATA 33691  
 QY 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeu 140  
 DB 33692 GACTATGCAGCCCGTGGTTGGTAGTGTGGCAGCATCTGGCAGGCATCTGCCAGTTC 33751

QY 141 LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu 160  
 DB 33752 CTGCGCTGGATGACCTCTCAAGTTGGTGAGCCCAACCATCGCAAGGAGCAGCATCTCG 33811  
 QY 161 SerGlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAsp 180  
 DB 33812 TCAGGGGCCAGGAGCCTGATTCCATGGCTGGCGCTCATTTGTTGAACAGGTCCAGGGGCGCT 33871  
 QY 181 AspLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSer 200  
 DB 33872 AGCAAGTTGTTGGGGCACCCTCTTGTATGCAAGTCTGAGTACTTCTCCACCTTGAAT 33931  
 QY 201 SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPhe----- 214  
 DB 33932 TCCCGTAAGCTCATCGACGATGAGCAGCAGTGAGATATGTTTCAGTCTGGTGGTGAGCTT 33991  
 QY 214 ----- 214  
 DB 33992 GTTGCTTTTGTCTTCTTCCCTGTTCTTATGAGCCAGGTCTTATTCACTCTTTC 34051  
 QY 214 ----- 214  
 DB 34052 TCAGTCTGGCCCTGCTTCTTAGGAAACTTTGGTTTGAGAGAGGAATTTATGACCCCTTC 34111  
 QY 214 ----- 214  
 DB 34112 ACTGTCACATAGATGTAGACCAGCTGTTGAGGACAGGAAGATGGCAGAGTTTACCTGTG 34171  
 QY 214 ----- 214  
 DB 34172 GAAAGTCTCTATAGAGAGGCTAAGTCACTTGCACCTCTGTGGCATGGAATCCAGTAA 34231  
 QY 214 ----- 214  
 DB 34232 TTGTCAGGTGGGCGAGATGGAGATCACTGCGGGAAGCCTTAGAAAAACCGCATTTG 34291  
 QY 215 -----Ala 215  
 DB 34292 AACAGGTTGGAGAGTGAAGTGGCCCTTACCCCTGCTCATACTTTGCTTAATCTCC 34351  
 QY 216 TyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThrPheThr 235  
 DB 34352 AG-GTGTACCAGGATGAGTTTGTCTCTTCAGATCAAGATCCCTTCAGACACATTTGCTCC 34410  
 QY 236 IleIleProAspPheAspIleTyrValTyrGlyPheSerSerGlyAsnPheValTyr 255  
 DB 34411 TTGTACCTCGCTTGTGACATCTACTACATATATGCTTTGTGTCAGTCCCTCTCTGCTGATC 34470  
 QY 256 PheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGluGlnVal 275  
 DB 34471 TTTCTAACGTTGACGCTGGACACCCAGCAGACGCTGTTGGATACAGCAGCGAGAAATTC 34530  
 QY 276 TyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGlu 295  
 DB 34531 TTCACATCCAGATGCTGCGCATGTGTCAGGGGACTCAGAGTTCTACTCTTATGTAGAG 34590  
 QY 296 ValProIleGlyCysGluArgSerGlyValGluTyrArgLeuGluAlaIleTyrLeu 315  
 DB 34591 TTCCCTATTGGCTGCTCTCGCGTGGTGGAGTACCGCTTGGTCAGAGAGTCCCATCTG 34650  
 QY 316 SerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValIleHisProAspAspLeuLeu 335  
 DB 34651 GCCAAGCCAGGCTGCTATGGCCCAAGCTCTGGGTGTGCCAGCGCAGGATGCTTC 34710  
 QY 336 PheThrValPheSerLysGlyGlnLysArgLysMetLysSerLysLeuAspGluSerAlaLeu 355  
 DB 34711 TTCACATCTTCTCTCAGGGCCAGAAACAGAGCCCAACCCACCTCGCGCAGACCATCTT 34770  
 QY 356 CysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyr 375  
 DB 34771 TGCCTTTTCAACCTCAGCAGCATTAATGCCCAACATCCCGGCCGCGAATCCAACTGCTGCTAC 34830  
 QY 376 ArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIlePro----- 392



```
QY 630 -----HisValValGlnLeuGlnLeuLysSerLysGluThrG1 643
Db 37046 CTGGCTTGTCAGAGGCCACTCATACTGTCGGCCTTCACCTGCTCCATGGAGACTGG 37105
QY 643 yMetThrPheAlaSerThrSerPheValPheTyrAsnCysSer----- 657
Db 37106 TGTGAGGTTTGTCTGGGGTTGACTTTGTCTTCTACAAGTCGAGTGCCTCCAGTCGTGAGT 37165
QY 657 ----- 657
Db 37166 ACCTGGTCCAGAAATCTGTCCCTAGCCCTGGGGATAGTAGATAAGTGTCTGAAATATATA 37225
QY 657 ----- 657
Db 37226 AAAATTGCACTGTAACCTTTCTCTAGGTTTCAGGGGTGAGTTGTTGGGACTTCTCGCTGTC 37285
QY 658 -----Va 658
Db 37286 GGCACCTTCTGTAGTCTCTAGTGCCTATATGTGGCTTGGCAGGTTCTCGCTACCCCTAT 37345
QY 658 lHis----- 659
Db 37346 TCATTGTTCCCAATGTCTTGTGGGTGGGTGATACAGGTGCGTATGACCTCTGTAT 37405
QY 660 -----AsnSerCysLysSerCysValGluSerProTyrArgCysHI 673
Db 37406 TTCTAAAGCTACTCCCCCTCCAGGTGTATGCTCTGTGTTGGCAGCCCTTACCCCTGCCA 37465
QY 673 sTrpCysLysTyrArgHisValCysThrHisAspProLysThrCysSerPheGlnGluG1 693
Db 37466 CTGGTGTAAAGTACCGCATGTGTGTACCAAGCCACCCACAGCAGTGTCTTTTCCAGGAGG 37525
QY 693 yArgValLysLeuProGluAsp----- 700
Db 37526 CAGGGTCCACAGCCCTGAGGT-CAGGTGGTGCAGTGCACCGTGGAGAGGCTGGGCTCCT 37584
QY 700 ----- 700
Db 37585 TGGGGGTGGCCCTCTCACATTTTGTGTTGACCTCCCATTTCCATGGCCCTGGAGAAGGC 37644
QY 701 -----CysProGlnLeuLeuAr 706
Db 37645 GGTAGACCAGTGTGCTCTAAACAAGTGTCTCTCCCGAGGCTGCCCTGAGATPCTGCC 37704
QY 706 gValAspLysIleLeuValProValGluValIleLysProIleThrLeuLysAlaLysAs 726
Db 37705 TCAAGGGGACCTCTTGATTCCTGTGGGTGTCATGCAGCCTCTAACTCTGGGGCTAAGAA 37764
QY 726 nLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGlySe 746
Db 37765 CTGTCCACAGCCCTCAGTCCCGCAGAGAAGAAATTATGAATGCGTGGTCCGTGTACAGGAGC 37824
QY 746 rGluGlnArgValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThrSe 766
Db 37825 GCAACATCGGGTACCTGCAGTGGCTTTAAACAGCAGCAGCGTGCAGTGCCAGAAATGCTC 37884
QY 766 r----- 766
Db 37885 GGTAAAGACCAGAGTGCAGGGTGGGCCAGCCAGGAATAATTTGTTTGTGTTGTTT 37944
QY 766 ----- 766
Db 37945 GTTTTATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 38004
QY 766 ----- 766
Db 38005 TTTGTATTTGGTGGTTTGGTCTATGCCCTAGCTTACTTATCTGGAACACCTTTTTTTTT 38064
QY 766 ----- 766
Db 38065 TTTTTTTTTTTTTTTTTTATAAAGGATTCCTAAAGTGGCGAAGGATTTTGTCTCATCTT 38124
QY 767 -----TyrSerTyrGluGlyMetGluIleAsnAs 776
Db 38125 ATTTGTGCTCATGTGTTCCATTTTCTTAGTACTTCTATGAAGGTGATGAGTTTGGTGA 38184
QY 776 nLeuProValGluLeuThrValTrpAsnGlyHisPheAsnIleAspAsnPro----- 794
Db 38185 CACCGAACTGGACTTCTCTGTGGTTTGGGATGAGATTTTCCCATTTGTAAGCCTCTCTAG 38244
QY 794 ----- 794
Db 38245 CTTCCGAGGTCAGGGGTGCATGCGGCCAAGCAGCATGCTCTGGGGCAGGATTCGCCAC 38304
QY 795 -----AlaGlnAsnLysValHI 800
Db 38305 TGCTGGGGCAAGTCTGTAGGTCTCTAGAGCCCTCTCTTTCTGCCAT-TGCCTGTCTCCA 38363
QY 800 s-----LeuTyrLysCysGlyAlaIleMetArgGluSerCy 811
Db 38364 TCAGTAGTCTCTGTGCTGTGCAGCCCTTCTTTACAAGTCTCTGGGCTCAGCGCCTAGCTG 38423
QY 811 sGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTyrCysGlnGlyProGlyG1 831
Db 38424 TGGCCTCTGCCTCAAGGCTGATCCCGATTCAACTGTGGCTGGTGCATCTCAGACACAG 38483
QY 831 nCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuGluLeuSerGlyAl 851
Db 38484 GTGCCAGTGGGGCTCAGTGCACAGTCCCAAGTCCCAAGTAACTGGATGCACCAAGCCAGAA 38543
QY 851 aLysSerLysCysThrAsnProArgIleThrGluIle-Ile----- 864
Db 38544 GGGTCCCGATGCAGCCATCCCGGAATCACCCAGGTTAGCTTCACTACAAATCTTACCT 38603
QY 864 ----- 864
Db 38604 TGCTGTTAGGATGTGTGAGTCTGGCCGTGATCTTCTTCTCTCTCTCTCTCCAGA 38663
QY 865 -----ProValThrGlyProArgGlyGlyThrLysValThrIleArgGlyGluAsnL 883
Db 38664 TTCACTCCACTTACAGGACCAAGAGGTGGCACCCGGCTCACCATTGTGGGTGAGAACC 38723
QY 883 euGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyValGluCysSerP 903
Db 38724 TGGGCTCTCATTCCCGTGGGTGAGTGGC-----CTTCGAGTAGTGTGTGACGTGTGCACT 38777
QY 903 roLeuValAspGlyTyrIleProAlaGlu-----GlnIleValCysGluMetGlyG 920
Db 38778 CCATTCCCAACCAATATGTCACTGTGTAAGAGTGAGTGTGGCTATCTGAGCACCCCTGTCC 38837
QY 920 luAlaA---LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValLysArgP 939
Db 38838 TTGTTTACCACAGCTGGCTGACCTTCTGTGACTCTTAGGATCGTATGTGAGATGGAG 38897
QY 939 roGluPheMetAlaArgSerSerGlnLeuTyrTyrPheMet----- 952
Db 38898 AATCGCTGGTGGCCACCCACCCCTGGCCCTGCTGAGCTCTGTGTAGGGGATGTGTCTG 38957
QY 952 CTGACTTCCGCACACAGTCCCAAGCACTCTACAGCTTTCTGGTGGCTGCTTAACCTAGCC 39017
QY 952 ----- 952
Db 39018 CTTCCCTTCCCTATCTGGGAAATGTGGAGTGTCCCTTGGATCAGCTGGCTTCTTACC 39077
QY 953 -----ThrLeuThrLeuSerAspLeuLysProSerArg 963
Db 39078 TCCTCTCAAGGGCTGCTTCTTCTCAGACCCCTACATTTGACCGTGTGAGTGTCTCTCGG 39137
QY 964 GlyPrometSerGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySer 983
Db 39138 GCCTCAGCTTCTGGAGGCACTCGGCTTACCATCTCTGGAAATTTCTCTGGATGCCGCGCAGC 39197
QY 984 AsnValValIleMetPheGlyLysGlnProCysLeuPheHis----- 997
```

Db	39198	AGGTCACAGTATATAGAGATGGCGAGTGCCAGTTGT- GAGGTGGCGCTGAGCACA	39256	Db	40276	GTGTGGTTATCCACTACTAGGGCAAGAACTGATCCCTGCTGCAGCTGGCAGCTCC	40332
Qy	997	-----	997	Qy	1178	LysLeuAsnThrValLeuValGlyGluLysProCysThrValThrValSerAspVal	1197
Db	39257	AGCCATCTGGAGTTGAGCACTGATTGTCTGAAGACTATGGGGCAGGGTAGGTCAAGGT	39316	Db	40333	CGCTCACTACACAGTTGATTGGAGCAGCCATGTGCACCTAATCTGTCTCAGATACT	40392
Qy	998	-----ArgArgSerProSerTyrrile	1004	Qy	1198	GlnLeuLeuCysGluSerProAsnLeuIleGlyArgHisLysValMet	1213
Db	39317	GCAGGCCATCTGCTCATGCCTACCTGCTTACCTGTGTAGGAGAGATGCCGAGGCATC	39376	Db	40393	CAACTTCTGTGTATCCCAAGCCAGACAGCGCGCAGCTCTTATGTGGTGTATGGC	40452
Qy	1005	ValCys---AsnThrThrSerSerAspGluValLeuGluMetLysValSerValcInVal	1023	Qy	1213	-----	1213
Db	39377	GTGTGATCTACCTGCTCCACCTGGGCCAGTCAGACCCCTATCACCTTGGCCATC	39436	Db	40453	TCAAGAGCCACACGGCAGCCCTGGCTGTGCATGGAGTTGGTTCACTCAAGCTTGCC	40512
Qy	1024	AspArgAlaLysIle---HisGlnAspLeuValPheGlnThrValGluAspProThrIle	1042	Qy	1214	-----AlaArg-ValGlyGlyMetGlnThrSerProGlyMetValThrIleAlaProAs	1231
Db	39437	GATCATGCCAACATCTCCACACTGGAGTCTATACCTACACCCAGGACCCCTACGGTC	39496	Db	40513	CTTACAGGTGCTGGGTGGCTGGAGTCTGTGGTGGCAGCCTGCACATCACTGCTGA	40572
Qy	1043	ValArgIleGluProGluTyrSerIleVal	1052	Qy	1231	pSerProLeuSerLeuProAlaIleValSerIleAlaValAlaGlyGlyLeuLeuIle	1251
Db	39497	ACACACCTTGAGCCCACTGGAGCATCAT-CAAGTGAGACCCCTGAGGAGATGGGTTGA	39555	Db	40573	TCGGGCACCTGACCTTACCAGCTATGTGGGGCTAGCAGCAGAGGGGGGGCTCTTGTCT	40632
Qy	1052	-----	1052	Qy	1251	ePheIleValAlaValLeuIleAlaTyrLysArgLysSerArgGluSerAspLeuThrLe	1271
Db	39556	GGACTGAGTTGGGCTCTGAGAGTCTTGGGGCTTACCTATTACTTGGGGCTTGGGTCAT	39615	Db	40633	GGCCATCACTGTGTCTGCTCGCTTACCAAGCTAAGCTACAGCTCAGGATGCACCCGACCT	40692
Qy	1052	-----	1052	Qy	1271	uLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAlaLeuGluCysLysG	1291
Db	39616	CAAGTGTCTCTCTCTCTGAGCCCTCATCAGTAGGGGCCACAGAGTTGTCTCTGCCCAT	39675	Db	40693	TAAGGGCTTTCAGTACAAATGGACAACCTGGAGTCTCTGTGGCTTGGAGTCAAGGA	40752
Qy	1052	-----	1052	Qy	1291	u-----	1291
Db	39676	TAGAGCTGGCACCTGCTGATGATGACCAAGGTACAGAGCTCTGAGCTCTCAGAAATCCT	39735	Db	40753	AGGTGCTTAAAGCAGGGGGTGTACCACATTGTGTGAGAGCTGAGGGACTTTCTCTCTCT	40812
Qy	1053	-----SerGly	1054	Qy	1292	-----AlaPheAlaGluLeuGlnThrAspIleHi	1301
Db	39736	TGGTAACCTTTGGACAGGGTGGGAGGAGTCTGAGCTCTTTACCCATCTCTGCAAGTGA	39795	Db	40813	TGTTCAACTCACTCTACCGTATTGTTTCCAGCCTTGTGAGCTGAGCTGATATCAA	40872
Qy	1055	AsnThrProIleAlaValTyrGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArg	1074	Qy	1301	sGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspTyrArgThrIle	1321
Db	39796	AGCACTCCATCACTGAGTGGAGAACCATCTGTGACAGTTCAGAACCCGCTGACGA	39855	Db	40873	TGAACGTACAAACCATGATGGTGTCCAGATCCCTTCTTGACTACCCGACCTATGC	40932
Qy	1075	AlaLysHisGlyLysGluHisIleAsnIle	1085	Qy	1321	rMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLeuGluVal-P	1341
Db	39856	GCCAAGTACCGTGTATTGAGACCACTAACCTGAGTACTGGGTCTTCTACCCCTGTTCTAC	39915	Db	40933	TGTGGCGTGTCTTCCCTGGCATCGAGCTCACCCAGTCTCAAGGAGCTGGATGTGAG	40992
Qy	1086	-----CysGlu	1087	Qy	1341	roGlyTyrArgGlnGlu-----	1346
Db	39916	CATTGCTACTCATGGCTTCATGGCTGAGATGATTGCTGTTCTTCACAGACATGCCAG	39975	Db	40993	CCTCTTACCCCAAGCCCTCCCATATTCTGTTTCCCCAGAGCCACCCTACTACTCT	41052
Qy	1088	ValLeuAsnAlaThrGluMetThrCysGlnAlaProAlaLeuAlaLeuGlyProAspHis	1107	Qy	1347	-----ArgValGluLysGlyLeuLysLeuPheA	1356
Db	39976	GTGATCAGCAGACTGCAATCTATGTAAAGCCCTGGCATCTTCTTGGGACCCCTCAG	40035	Db	41053	CTGAGATTCCCTGGTCCCAACAGACTCCCCCAATGTGGAAGGCCCTCGCTATATTG	41112
Qy	1108	GlnSerAspLeuThrGluArgProGluGluPheGlyPheIleLeuAspAsnValGlnSer	1127	Qy	1356	lAsnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnA	1376
Db	40036	CCTCGGGCCCAAGCGGACCCCTCAGCAGTTTGGCTTCTTGTGGACACCGCTGAGGCA	40095	Db	41113	GACAGTTGTGCACAGCCGGCCCTCTCTGCTCACCTTATCCACACTTTGGAGGCCGAGA	41172
Qy	1128	LeuLeuIleLeuAsnLysThrAsnPheThrTyrTyrProAsnProValPheGluAlaPhe	1147	Qy	1376	rgSerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnS	1396
Db	40096	GCCCGCTCCCTCAACCCGTTCTTCTTCCCTACTACTACCTGATCCCGCTTTGAACCATTT	40155	Db	41173	GTAGTTTCTCCATGCGTGACCGTGGTACTGTGSCCTCACTCACCATGGTGGCCCTGCAGA	41232
Qy	1148	GlyProSerGlyIleLeuLeuLysProGlyThrProIleIleLeuLys	1164	Qy	1396	erLysLeuGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysA	1416
Db	40156	GGCCCTCTGGTGTAGATCTCAACCTGTTTACATGTTGTTATGAAGTGTGGGG	40215	Db	41233	GGCGGCTTACTATGCATCGGCTGGCTGTCTCAAGCAACTGTGGTGTGACTCATAGAGAAA	41292
Qy	1164	-----	1164	Qy	1416	snLeuGluSerLysAsnHisProLysLeuLeuLeuArg-----	1428
Db	40216	ATAGGATAGGAGTAGAAGAGAGGAGGAGAACTGAGAGCTGGAGCTGTGCTTGTCT	40275	Db	41293	ACCTTGAGACAGAACCCCAAGCTGTATTGCCAGGTATGTAACTGTCTTGTCTGTTG	41352
Qy	1165	-----GlyLysAsnLeuIleProValAlaGlyAsnVal	1177	Qy	1428	-----	1428
				Db	41353	GCCTCTCTCCAGTCTCTCTCTCCATTCTCTGCTTGGGGCCCTACCTCTGTGTCAAGTCTC	41412

QY 1429 -----ArgThrGluSerValAlaG 1435  
Db 41413 ACCCTTGTGAGCCAGGGCTTAGGACCTATCTCTTACCATAGGACAGAGTCGGTGGCTG 41472  
QY 1435 LuLysMetLeuThrAsnTrpPheThrPheLeuLeuTrpLysPheLeuLys----- 1451  
Db 41473 AGAAGATGCTACCAACTGGTTTACATCTCTGTCATAGTTTCCTGAAGGTGTGCTGGG 41532  
QY 1451 ----- 1451  
Db 41533 GTGGGCTGGTGCAGGAGAGGAATGTCGGGGCTGGGAGCAATTCATTGATTAAGCCAGA 41592  
QY 1452 -----GluCysAlaGlyGluProLeuPheS 1460  
Db 41593 CAGGCAGCCAGCTATGCCCTTAACCTGGGTGTAGGAGTGGCGGGGAGCCACTCTTCC 41652  
QY 1460 erLeuPheCysAlaIleLysGlnMetGluLysGlyProIleAspAlaIleThrGlyG 1480  
Db 41653 TGCTGTACTGCGCCATCAAGCAGCAGATGGAAGAAGTCTTATGTATGCCATTAACAGGCG 41712  
QY 1480 luAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLysThrL 1500  
Db 41713 AGGCCGCTACTCTTAAGTGAGGACAGCTATCCGGCAGCAGATCGACTATTAAGACC 41772  
QY 1500 euVal----- 1501  
Db 41773 TGGT-GAGCATGGGACTATCGCAGCAGGAAGTAAGGTGGAGCCTTGCTTAGGAATCATGC 41831  
QY 1502 -----LeuSerCysValSerProAspAsnAlaAsnSerProG 1514  
Db 41832 TGAGGTTCCTTCACTTCACTTCACTTGCATGTGTGCGCGGAGGAGGCGAGTGC 41891  
QY 1514 luValProValLysIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuA 1534  
Db 41892 AGGTCCCTGTGAAGGTTCCTTAAGTGTATGATCATCCAGCCCAAGACAAGCTGTGG 41951  
QY 1534 spAlaIlePheLysAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeu- 1553  
Db 41952 ATACTGTGTACAAGGGTATTCATACTCTCAGCGCCCAAAAGCTGAGGACATGGATT-G 42010  
QY 1553 ----- 1553  
Db 42011 GGTAAGTCCCTCCCTTTCTCCTAGCTCCCACTATCCAACTCCCTCCAGGAGGTCTC 42070  
QY 1553 ----- 1553  
Db 42071 ATATTCTGTATCTGAGCTCTGTCTTAAAGGAATGTCTAAGTCTTGTGAGGGTGA 42130  
QY 1554 -----GluTrpArgGlnGlySerGlyAlaArgMetI 1564  
Db 42131 ATTTCCTTAAGAGCTGACTATGGCTTGCAGAAATGGCGCCAGGCGCGCATGGCCCGAATCA 42190  
QY 1564 leLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTrpLysArgLeuAsnThrL 1584  
Db 42191 TCTCAAGATGAGGACATCACTACAAAGATTGAGTGTGACTGGAAGAGGGTCAACTCAT 42250  
QY 1584 euAlaHisTyr----- 1587  
Db 42251 TGGCCCACTACCAGGTGAGGAATTTGAACCTATTCAAGTCCAGGAGCACTAGAAGTAG 42310  
QY 1587 ----- 1587  
Db 42311 TGGCCAGGTCTCTCTCGTCCCTTATCCCTATATCTATGAGGCTTCTCATAGGGCCTTG 42370  
QY 1587 ----- 1587  
Db 42371 ATTGCCACTGTACATATATATACTGAAGAGATGAATGAATGTGTGCTCCCTGGCTCCTT 42430  
QY 1588 -----GlnValProAspGlySerValValAlaLeuValSerLysGlnV 1602  
Db 42431 CTGTTTCTTTGGACCTCCAGGAGTGTGTTCTTTAGTACACTGGTGCCCAACAAAG 42490

QY 1602 alThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSer----- 1617  
Db 42491 TGTCTGCTATAACATAGCCCAACTCGTTCACCTTCACCCGGTC-ACTTAGTCTGCTACGCT 42549  
QY 1617 ----- 1617  
Db 42550 AGGTATCCATATGTGTGGCTTTTCTCTGGTCCCGTGGCCTTTAATAAGGCTGTATCTTA 42609  
QY 1617 ----- 1617  
Db 42610 GGAATCCCCAAGTAGCTGTCTTGTATCCATTGGAAGACTTTCATTATTGTTGGGCCCTG 42669  
QY 1617 ----- 1617  
Db 42670 GGTGGTCTTCAGTTGTATAGAGCCCTAATAGGAGGGGTGGCAAGGATAGGTGGTCCCCAC 42729  
QY 1617 ----- 1617  
Db 42730 AAAGACTTCTAGAAGATAAATGGACAAGAGGTGGGTAAATGCCAGAGCCTCTAGGGCTT 42789  
QY 1618 -----AlaSerLysTyrGluAsnMetIleArgTyrThrGlySerProAspS 1633  
Db 42790 AAGAGTCACATCTCTCTCTCCCTAGAGAGCTTGTCTCGTCTGCCAGCAGCCCGGATA 42849  
QY 1633 erLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyValLysMetTrpH- 1653  
Db 42850 GCCTCCGTTCCGAGCACCTATGTCTACGCTGACCGAGGAGGACGACCAAGCTGTGC 42909  
QY 1653 lsLeuValLysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetV 1673  
Db 42910 ACCTGGTGAGGAACACACCCACTGTATCCAGAGAGGAGACCGCGCAGCAAGATGG 42969  
QY 1673 alSerGluIleTyrLeuThrArgLeuLeuAlaThrLys----- 1685  
Db 42970 TCTCAGAAATATATCTCACAAGGTGTCTGCCACCAA-GGTGTGGGCTGCCCATAGAC 43028  
QY 1685 ----- 1685  
Db 43029 ACCTTCTAGCCAGCTCTTAAACATGAGATGAGTGTCTTCTTAGGGCCTTGGCAGTAGGA 43088  
QY 1685 ----- 1685  
Db 43089 CCTCTACGCTGAGGTCAAGCAAGGTCTATTTTAGGAATGTTGGCCTTTCAGAGAAGTT 43148  
QY 1685 ----- 1685  
Db 43149 TTTTCTTACTGAGCTATTTCTGCCCTCTCTCTCAGATTTTGGGAACCTTCTGTACATGCT 43208  
QY 1686 -----GlyThrLeuGlnLysPheValAspAspLeuPheGluThrI 1699  
Db 43209 TCCTTTTACCATCACTCATAGGCGACGTTGCAGAGTTTGTAGATGACTTGTTTGAACATG 43268  
QY 1699 lePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspP 1719  
Db 43269 TGTTCAGTACAGCCACCGGGCTTACCTTACCTTGGCCATCAAGTACATGTTTGACT 43328  
QY 1719 heLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrpL 1739  
Db 43329 TCCTGGATGAACAGGCTGACCAGCCAGATCAGTGACCTGATGTGGCTCACACCTGGA 43388  
QY 1739 ysSerAsn----- 1741  
Db 43389 AGAGCAACTGCTACTTACTGGTTTCTGTACTGTCTGTGGCTTCTCTGAGGGTGGTGGAG 43448  
QY 1741 ----- 1741  
Db 43449 GAAGAGGCTAGTTACTTATGACCAAGCCTCCAGGCCAATGGAGACAGGTTGACTAGCAACTG 43508  
QY 1741 ----- 1741  
Db 43509 AATAGCCTTCCATGCAGGAGGATGGACAGTGGTATGGCCTGGGAAGTCAATATGAGTAT 43568  
QY 1741 ----- 1741

43569	ATAGAAGCTGGTGCTGCTGGTTGTGCAGATGGGAAGTGTCTAGTCTAGCACTGGGATTG	43628			
1741	-----	1741			
43629	ACCATTTTCTCTAGTGTCTGTGTCAGGGTAGACACACTGATGGCTTGGCATAGTGGAG	43688			
1741	-----	1741			
43689	ACAGTGAAGTTTCAGGGGTGAGGTTGAGGACAGAAAGCCATCTATTCTTATTCAT	43748			
1741	-----	1741			
43749	TGCCTGAGTCGACGAGAAAGTTAAAGAAAACCCATTATCCAAATAGTCTCTCCTGTT	43808			
1741	-----	1741			
43809	ACTTCAGGACAAGATATGACCTGCGCCAAACAGGACAAAGGTGGAAGGACGAGTGGGG	43868			
1742	-----CysLeuPro	1744			
43869	AAGGGGGCACATAGGGCCATTAGGAGGCTCAAGTGTATGTCCCTCTCTTCAGCTTACCT	43928			
1745	LeuArgPheTrpValAsnMetIleLysAsnProGlnPheValPheAspIleHisLysAsn	1764			
43929	CTGCGTTTCTGGTGAATGTGATCAAGAATCCGCAATTTGTGTGACATCCATAAGAA	43988			
1765	SerIleThrAspAlaCysLeuSerValValAlaGlnThrPheMetAspSerCysSerThr	1784			
43989	AGCATCACAGCGCTGTTTGTCACTGTGTGGCCAGACCTTCATGGAGTCCGTGTCTACA	44048			
1785	SerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyAlaLysAspIle	1804			
44049	TCAGAGCACCGCTGGGCAAGGATCCTTCCACAAAGCTCCCTATATGCCAAGGATATT	44108			
1805	ProSerTyLysAsnTrpValGlu-----	1812			
44109	CCCAATTACAAGAGCTGGGTGGAGAGTGGGCCCTGGGATACTGCTGTGTATGCTGTG	44168			
1812	-----	1812			
44169	GCAGGAGGAAGTGCCTAAACAGCGCTGTCTGAGGCTGGCCAGATACAAAGAAAGCCCTCA	44228			
1812	-----	1812			
44229	GAGGCTGGCCAAACACTGATGCATGAGTGTGGCATTTGTGACCTGTGCATGATTCACT	44288			
1813	-----ArgTyTySerAspIleGlyLysMetProAlaIleSerAspG1	1827			
44289	CACCCAAAGTCTCTATAGGTACTATCGAGATATCCGAAAGATGGCATCCATCAGTGACCA	44348			
1827	nAspMetAsnAlaTyLysLeuAlaGluInSerArgMetHisMetAsnGluPheAsnThrMe	1847			
44349	GGCATGGAGCGCTACTTAGTGGAGCACTCCGCTCCATGCTATGACTTCAATGCTCT	44408			
1847	tSerAlaLeuSerGluIlePheSerTyValGlyLysTySerGluGluIle	1864			
44409	AAGTGCACCTCAGCGAGCTCTACTTCTATGCTACCAAGTACCCTCAGGAGGTA	44460			
RESULT 12	AL807376/c	204152 bp	DNA	linear	HTG 13-AUG-2002
LOCUS	Mus musculus chromosome X clone RP23-436K3, *** SEQUENCING IN				
DEFINITION	PROGRESS:***, 3 unordered pieces.				
ACCESSION	AL807376				
VERSION	AL807376.2				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
	1 (bases 1 to 204152)				
	Ellwood,M				

Direct Submission			
Submitted (10-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,			
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:			
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk			
On Aug 15, 2002 this sequence version replaced gi:21628092.			
----- Genome Center			
Center: Wellcome Trust Sanger Institute			
Center code: SC			
Web site: http://www.sanger.ac.uk			
Contact: humquery@sanger.ac.uk			
----- Project Information			
Center project name: BM436K3			
----- Summary Statistics			
Assembly program: XGAP4; version 4.5			
Chemistry: Dye-terminator; 100% of reads			
Consensus quality: 203585 bases at least Q40			
Consensus quality: 203744 bases at least Q30			
Consensus quality: 203854 bases at least Q20			
Insert size: 203952; sum-of-ctnigs			
Insert size: 205924; 5.1% error; agarose-fp			
Quality coverage: 8.31x in Q20 bases; sum-of-ctnigs Quality			
coverage: 8.28x in Q20 bases; agarose-fp			
-----			
* NOTE: This is a 'working draft' sequence. It currently			
* consists of 3 ctnigs. The true order of the pieces			
* is not known and their order in this sequence record is			
* arbitrary. Gaps between the ctnigs are represented as			
* runs of N, but the exact sizes of the gaps are unknown.			
* This record will be updated with the finished sequence			
* as soon as it is available and the accession number will			
* be preserved.			
-----			
* 1 18988: contig of 18988 bp in length			
* 18989 19088: gap of 100 bp			
* 19089 158474: contig of 139386 bp in length			
* 158475 158374: gap of 100 bp			
* 158575 204152: contig of 45578 bp in length.			
-----			
FEATURES			
source			
1..204152			
/organism="Mus musculus"			
/db_xref="taxon:10090"			
/chromosome="X"			
/clone="RP23-436K3"			
/clone_lib="RPC1-23"			
misc_feature			
1..18988			
/note="assembly_fragment:00260			
fragment_chain:1			
clone_end:SP6			
vector_side:left			
misc_feature			
19089..158474			
/note="assembly_fragment:01548			
fragment_chain:1			
misc_feature			
158575..204152			
/note="assembly_fragment:01957"			
BASE COUNT			
55626 a 47968 c 46690 g 53668 t 200 others			
ORIGIN			
Alignment Scores:			
Pred. No.: 0			
Score: 4403.00			
Length: 204152			
Matches: 1115			
Percent Similarity: 37.79%			
Conservative: 292			
Best Local Similarity: 29.95%			
Mismatches: 431			
Query Match: 44.07%			
Indels: 1897			
DB: 2			
Gaps: 43			
-----			
US-09-964-956-13 (1-1896) x AL807376 (1-204152)			
Qy	8	TriThrCysLeuLeuSerHisLeuLeu-----MetValGly-----Met	20
Db	90264	TGTCATGCCACTGCTGTC-CTTCTCCATTCCTTCACCATAGGAGGCTGCTG	90206
Qy	21	GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe	40
Db	90205	GGTAGCAGC-----AGGCCATT	90188



Db	88032	GAATGTGAACATGAGCCTGGATAACCCCTGAATAGGGTGTAAATAACTTCCCAAGTTAGGCT	87973	Db	86955	GGTGGGGTCCGTCCCAATAATGTGTCACTGACATCCTCTGGGGTGCA-CGGTAAGTGACA	86897
QY	457	-----	457	QY	576	-----	576
Db	87972	CCAAGAAGTGAATCCTGTTCTTTGGTGGCTCTTGCCATCTACCTTCGCCAGCATTTGT	87913	Db	86896	TAGAGATACATTGGCCCTGTATATGCTTCATGAGGAGGGTGACACCTGCTCTGAACCCCT	86837
QY	457	-----	457	QY	577	-----	592
Db	87912	TCGAAGTAGCATTTGCTTGTAGCTTAAGGCTATGGCTGTGGAAGTAAAGCCCTACAAC	87853	Db	86836	CTGCCCTCCCAAGCTGACCGTAGCCATAGCCATGGCCAAAGTCCAGACCTCAGTGTGGGTGAG	86777
QY	457	-----	457	QY	592	nCysThrPheGluAspLeuSerGluMetAspGlyLeuValVal- -GlyAsnGlnIleGln	611
Db	87852	TGAATACAGACTCAAGAGGCGCATATCTGACCAGCTCAACTGTAATCTAGAAAGTCCCTG	87793	Db	86776	CTGTTCTTCTTGGAGGCTGACTGAAGTGGGCTATTCTGCTGCCCTCTCGAAGATTACG	86717
QY	457	-----	457	QY	611	nCysThrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAsp- -	629
Db	87792	GCCGAAGAGCCTGCTCTGGGTGCAGGAGCTATGGTAGCATGAAGCTGGCTGTGCATTTT	87733	Db	86716	ATGCTCTTCCACCATCCCTTCAGGAGCTCCAGACACTTACCAGAGGCGATGGTCAGTGGGA	86657
QY	458	-----	469	QY	629	-----	629
Db	87732	CTCTGCTCTGCTGTGGCTCAGGTGGGTGATGGCTGCCAGTGCCTCAG- -GATGCCAGCTA	87676	Db	86656	TGGGGCAATCTGGGGTGTGTTGGGGCAGGCGCAGANTTCCAGACTAGTTGGCAATAGCCAT	86597
QY	470	TyrGluThrValGlnValValAspProGlyProValLeuArgAspMetAlaPheSerLys	489	QY	630	-----	643
Db	87675	TATGAACCGTCTCTGTGGTCAGGCGCAGCCCATACTTCGAGACCTGCTCTCAGCCCT	87616	Db	86596	CTGGCTTGTGTCAGGGGCCACTCATCTACTGTGGGCTTCAGCTGCTCTCCATCGGAGATGG	86537
QY	490	AspHisGluGlnLeuTyrIleMetSerGluArgGlnLeu-ThrArg- -	504	QY	643	yMetThrPheAlaSerThrSerPheValPheTyrAsnCysSer- -	657
Db	87615	GACCACCACACATCTACCTCTGAGTGAGAACAGAGGTGGGCCCTGGTGGTGGTG	87556	Db	86536	TGTGAGGTTTGTCTGGGTTGACTTTGCTTCTTCTACAACTGCAGTGCCTCCAGTCGTGAGT	86477
QY	505	-----	506	QY	657	-----	657
Db	87555	TGTGTGGGCAAGGTGGGTGTGCTGATGTGGCTGTCCCGAGGTGAGCCAACTCC	87496	Db	86476	ACCTGTCCAGAAATCTGCCCTAGCCCTGGGGATAGTAGATAAGTCTCTGAAATATATA	86417
QY	506	ovaGluSerCysGlyGlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHi	526	QY	657	-----	657
Db	87495	GGTGAGACCTGTGAGCAGTACTCAGCTGTGCTGATGCTGGGCTCAGGGAGCCACACA	87436	Db	86416	AAATTTGCACTGTAACCTTTCTCTAGGTTTCAGGGGTGAGTTGTTGGAGCTTCTCTCTC	86357
QY	526	sCysGlyTrpCysValLeuHisAsn- -	534	QY	658	-----	658
Db	87435	CTGTGGTGGTGTGCTACACACAGGTGAGGAAGAGGGTGTGGGATGTGGGCGACGA	87376	Db	86356	GCACCTTCTGTCTAGTCTGATGTCCTATGTGGTTGGGAGGTCTTGCTACCCCTAT	86297
QY	534	-----	534	QY	658	lHis- -	659
Db	87375	GTGGCTCCCTGGGCCATTATCTTGGGACAGAGAGATTAGGAGACCTTTCACAGCTG	87316	Db	86296	TCATTTGTTCCCAATGTGTCTTGTGGTGGGTGATACAGTCCGTATGACCTCTGTTAT	86237
QY	534	-----	534	QY	660	-----	673
Db	87315	GACAGGACGACACACTGCTATATCTTCTTCCACCACCTGCCAGTTTTCGCCAGC	87256	Db	86236	TTCTAAAGCTACTCCCCCTCCAGGTGTATCTCTGTGTTGGCAGCCCTTACCCCTGCCA	86177
QY	534	-----	534	QY	673	stPcCysLysTyrArgHisValCysThrHisAspProLysThrCysSerPheGlnGlnGln	693
Db	87255	CTTGAGCCTTCCCCCTCCCTCTCTCCACCTTCTCCAGCTAATGAGCGGAACCTCTAC	87196	Db	86176	CTGGTGAAGTACCGTGTATGTGTACCGCCACACAGAGTCTCTTTTCCAGGAGGG	86117
QY	534	-----	534	QY	693	YArgValLysLeuProGluAsp- -	700
Db	87195	CCTGAGTGACGTCTTGGATCTCAGGGGATGAGCCCAAGGTACACTGTTGCTGGCAGT	87136	Db	86116	CAGGTTCCAGCCCTCAGGT- -GAGTGGTGCAGTCGACCGTGGAAAGAGGTGGGCTCCT	86058
QY	534	-----	534	QY	700	-----	700
Db	87135	ATGAGCTGGGTATGCTGCTTCTGTCCACCTCAGCCTCTTACTGCACCCCATCTGGGC	87076	Db	86057	TGGGGGCTGGCCTCTCACATTTTGTGTTGACCTCCCATTTCCATGGGCTGGAGAAAGC	85998
QY	535	-----	540	QY	701	-----	706
Db	87075	CCTCTGGGACAGTCTCGGACTGCTCTGGCCTGTCTCCATGCAAGGTGCTGCGTGAAGG	87016	Db	85997	GGGTAGACCAAGTGTCTTAAAGTGTCTCTCCAGGCGCTGCCCTGAGATCTCTGCC	85938
QY	540	uArgCysGluArgSerLysGluProArgPheAlaSerGluMetLysGlnCysValar	560	QY	706	gValAspLysIleLeuValProValGluValIleLysProIleThrLeuLysAlaLysAla	726
Db	87015	GCCCTGTCAGGTGCTCTGCCCATCTGCTGGTGGTGGTGTATGAGTGTGACAGGAGT	86956	Db	85937	TCAAGGGACCTCTTGTATCTCTGTTGGGTGTATGAGCTCTTAACTCTGCGGCTAAAGAA	85878
QY	560	gLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeu- -	576	QY	726	nLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGlyse	746
		-----	576	Db	85877	CCTGCCACAGCCTCAGTCCCGACAGAGATTTATGAGTGGGTGGTGTACAGGAGC	85818



Db	83666	GGTGATCAACGACACTGCAATGCTATGTAAGGCCCTGGCATCTTCCTGGGCAACCTCA	83607	Db	82589	TCTGAGATTCCCTGGTCCCCACAGACTCCCCCAATGTGGAGAGCCCTGCGCTATT	82530
Qy	1107	sGlnSerAspLeuThrGluArgProGluGluPheGlyPheIleLeuAsnValGlnSe	1127	Qy	1356	AlaGlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGln	1375
Db	83606	GCCTGGGCCCAAGGACGACCTGACGAGTTTGGCTTCTGTGGACACAGTGCAGGC	83547	Db	82529	GGACAGTTGCTGCACAGCGCTCCCTGCTCCACCTTTATCCACACTTTTGGAGGCCAG	82470
Qy	1127	rLeuLeuIleLeuAsnLysThrAsnPheThrTyrrProAsnProValPheGluAlaPh	1147	Qy	1376	ArgSerPheSerMetArgAspArgGlyAsnValAlaSerIleIleMetThrValLeuGln	1395
Db	83546	AGCCCGCTCCCTCAACCGTTCTCCTCACCTACTACCTGATCCGACCTTTGAACCACT	83487	Db	82469	AGTAGTTTCTCCATCGTGACCGTGTACTGTGGCTCACTCACCATTGTGGCCCTGCAG	82410
Qy	1147	eGlyProSerGlyIleLeuGluLeuLysProGlyThrProIleIleLeuLys-----	1164	Qy	1396	SerLysLeuGluTyrrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLys	1415
Db	83486	TGGGCCCTCTGCTGTAGATGTAACACCTGGTTTCATGTTGATTGAAGGTCTGGG	83427	Db	82409	ACCCGGCTTGACTATGCCACTGGCGTGCACAACAACCTGCTGCTGACCTCATAGAAA	82350
Qy	1164	-----	1164	Qy	1416	AsnLeuGluSerLysAsnHisProLysLeuLeuLeuArg-----	1428
Db	83426	CATAAGGATAGGAGTAGAAGGAGGGAAGAACCGAGAGCTGGAGTGTGCTGACTTTC	83367	Db	82349	AACTTGACAGCAAGAACAACCAACCAAGCTGCTATTGCGCAGGTATGTAAGTCTT	82290
Qy	1165	-----GlyLysAsnLeuIleProProValAlaGlyGlyAsnVa	1177	Qy	1428	-----	1428
Db	83366	TGCTGGGTTATCCACTACTAGGCAAGAACCTGATCCCTGCTGCAGCTGGCAGTCC--	83309	Db	82289	GGCCTCTCCTCCAGTCTCTGCTTCCATTTCTGCTGGGGCCCTACCTCTGTTCAGTCT	82230
Qy	1177	lLysLeuAsnTyrrThrValLeuValGlyLysProCysThrValThrValSerAspVa	1197	Qy	1429	-----	1434
Db	83308	-CGCCTCAACTACACAGTGTGATTGGAGGACAGCATGTGCACCTAACTGTCTCAGATAC	83250	Db	82229	CACCCTTGTGAGCCAGGCGCTTAGGACCTATCTTTACCATAGGACAGAGTCGGTGGCT	82170
Qy	1197	IcLnLeuLeuCysGluSerProAsnLeuIleGlyArgHisLysValMet-----	1213	Qy	1435	GluLysMetLeuThrAsnTrpPheThrPheLeuLeuTyrrLysPheLeuLys-----	1451
Db	83249	TCAACTTCTGTGATTGCCAAGCCAGACAGCGCCGCGAGCCTGTTTGGTGGATGGG	83190	Db	82169	GAGAAGATGCTACCAACTGGTTTCACATCTCTGCTGCATAGTTCTCTGAAGGTGTGCTGG	82110
Qy	1213	-----	1213	Qy	1451	-----	1451
Db	83189	CTCAAGAGCCACAGGCGAGCCCTGGCTGCTGCATGGGAGTTGGTTCACTCAAGCTTG	83130	Db	82109	GGTGGGTGGTGCAGGAGAGGAAATGTCCGGGGGTGGGCAATTGATTGATAAGCCAG	82050
Qy	1214	-----AlaArg-ValGlyGlyMetGluTyrrSerProGlyMetValTyrrIleAlaProA	1231	Qy	1452	-----GluCysAlaGlyGlyProLeuPhe	1459
Db	83129	CCTTACAGGTGCTGGTGGCTGGCTGGAGTTCTGGTTGGCAGCGCTGCACATCACTGCTG	83070	Db	82049	ACAGGCAGCCAGCTATGCCCTAAACCTGCTGGTGTAGGAGTGCAGCGGGGAGCCACTCTTC	81990
Qy	1231	spSerProLeuSerLeuProAlaIleValSerIleAlaValAlaGlyGlyLeuLeuIleI	1251	Qy	1460	SerLeuPheCysAlaIleLysGlnGlnMetGluLysGlyProIleAspAlaIleThrGly	1479
Db	83069	ATCGGCACCTGACCTTACCACTGATGGTGGGTGACGACGAGCGGGCGGCTCTGCTGC	83010	Db	81989	CTGCTGACTGCGGCATCAACAGCAGATGGAGAAGTCTCTATTGATGCCATAACAGC	81930
Qy	1251	IlePheIleValAlaValLeuIleAlaTyrrLysArgLysSerArgGluSerAspLeuThrL	1271	Qy	1480	GluAlaArgTyrrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrrLysThr	1499
Db	83009	TGCCCATCACTGTTGCTGCTGCTACAAAGCGTAAGACTCAGATGTCAGACCGGCACAC	82950	Db	81929	GAGCCCGCTACTCTCTTAAGTGAGGACAAGCTCATCCGCGCAGCAGATGACTATAAGACC	81870
Qy	1271	eulysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAlaLeuCluCysLysG	1291	Qy	1500	LeuVal-----	1501
Db	82949	TTAAGCGGCTTCAGCTACAAATGACAACCTGGAGTCTCGTGTGGCTTGGAGTGCAGG	82890	Db	81869	CTGGT-GAGCATGGGACTATCGCAGCAGAACTAAGTGGAGCCCTTGTAGGAATCATG	81811
Qy	1291	lu-----	1291	Qy	1502	-----LeuSerCysValSerProAspAsnAlaAsnSerPro	1513
Db	82889	AAGTGCTTAAGCAGGCGGTGTACACATTGTGTGAGAGCTGAGGACCTTCTCGCTTC	82830	Db	81810	CTGAGTTCTTCACTCTTAGACTCTGCAATTCTGTGTGCCGCGAGAGCGAGCGCAGTGCC	81751
Qy	1292	-----AlaPheAlaGluLeuIleThrAspIleH	1301	Qy	1514	GluValProValLysIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeu	1533
Db	82829	TTGTTCAACTCACTCTACCTGTTGTTCCAGCCTTGTGCTGAGCTGCAGACTGATATCA	82770	Db	81750	CAGTCCCTGTGAAGGTTCTTAACCTGTGATGATCATCCAGCCGCAAGACACAGCTGTTG	81691
Qy	1301	isGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspTyrrArgThrTyrr	1321	Qy	1534	AspAlaIlePheLysAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeu	1553
Db	82769	ATGAACATGACAAACACATGATGGTGTCCAGATCCCTCTTCTGGACTACCGGACCTATG	82710	Db	81690	GATACTGTGTACAGGGTATTCCATCTCTCAGCGCCCCAAAGCTGAGGACATGGATT--	81632
Qy	1321	hrMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLeuGluVal-	1340	Qy	1553	-----	1553
Db	82709	CTGTGCGGTCTTCTCCCTGGCATCGAGGCTCACCCAGTGTCTCAAGAGGTGGATGTA	82650	Db	81631	GGCTAAGTCCCTCCCTTTTCTCTAGTCCCACCTATCAATCCCTCCAGGAGCTCT	81572
Qy	1341	ProGlyTyrrArgGlnGlu-----	1346	Qy	1553	-----	1553
Db	82649	GCCTCTACCCCCCAAGCCCTGCCCATATCTGTTTCCCAGAGCCACCCTACTACTCT	82590	Db	81571	CATATTCTGTATGTAGCTCTGTTCTTAAAGGAATGTCTCTTAAAGGAATGTCTTAAAGG	81512
Qy	1347	-----ArgValGluLysGlyLeuLysLeuPhe	1355	Qy	1554	-----GluTrpArgGlnGlySerGlyAlaArgMet	1563
				Db	81511	AATTTCCCTAAGAGCTGACTATGGCTGCAGAAATGGCGCCAGGGCCGCATGGCCCCGAATC	81452

QY 1564 IleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTrpLysArgLeuAsnThr 1583  
|||||  
Db 81451 ATCTCCAGATGAGGACATCACTACAAAGATTGAGTGTGACTGGAGAGGGTCAACTCA 81392  
QY 1584 LeuAlaHisThr----- 1587  
Db 81391 TTGGCCCACTACCAGGTGAGGAACCTTTGAACCTATTAGCTCCAGGAGACACCTAGAAAGTA 81332  
QY 1587 ----- 1587  
Db 81331 GTGGCCAGGTCTCTCGTCCCTTTATCCCTATACTCATGAGGCTTCTCATAGGGCCTT 81272  
QY 1587 ----- 1587  
Db 81271 GATTGCCACTGTACATATATATACTGAAGAGTATGAATGAATGGTGTCCCTGGCTCCT 81212  
QY 1588 -----GlnValProAspGlySerValValAlaLeuValSerLysGln 1601  
Db 81211 TCTGTTTCTTTGGACCTCCAGGTGACTGATGTTCTTTAGTAGCACTGGTGCCCAACAA 81152  
QY 1602 ValThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSer----- 1617  
Db 81151 GTGCTGCTTAAACATGCGCAACTCTTCACCTTCACCCGGTC- ACTTAGTCGCTACGG 81093  
QY 1617 ----- 1617  
Db 81092 TAGGTATCCATATTGTGGTGGCTTTCTCGTCCCGTGCCCTTTAAATTAAGGCTGTATCTT 81033  
QY 1617 ----- 1617  
Db 81032 AGGAATCCCCAAGTAGCTGTCTTGTATCTCAATTTGAAAGACTTTTCATTATTGCGCCCT 80973  
QY 1617 ----- 1617  
Db 80972 GGTGTCTTTCAGTTGTATAGACCTTAATAGGAGGGTGGCAAGATAGTGGTCCCA 80913  
QY 1617 ----- 1617  
Db 80912 CAAAGACTTCTAGAAGATAATGAGCAAGAGGTGGTAAATGCCAGAGCCTGTAGGGCT 80853  
QY 1618 -----AlaSerLysTyrGluAsnMetIleArgTyrThrGlySerProAsp 1632  
Db 80852 TAAGAGTCACATCTGCTCTCGCTAGAGAGGTGCTCGCTGCTGCCAGCAGCCGGAT 80793  
QY 1633 SerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyValLysMetTrp 1652  
Db 80792 AGCTCCGTTCCGAGCACCTATGCTACCCCTGACCAGGAGGAGGTACCAAGCTGTGG 80733  
QY 1653 HisLeuValLysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGlySerLysMet 1672  
Db 80732 CACCTGGTGAGGAACCAACACCACTGATCACCAGAGAGAGACCCGGCAGCAAGATG 80673  
QY 1673 ValSerGluIleTyrLeuThrArgLeuLeuAlaThrLys----- 1685  
Db 80672 GTCTCAGAAATATATCTCAGAGGCTGCTGCCACCAA- GGTGTGGGCTGCCCCATAGA 80614  
QY 1685 ----- 1685  
Db 80613 CACCTTCTAGCCAGTCTTAACATAGATGAGTGTCTTTCTTAGGGCCTTGGCAGCTAGG 80554  
QY 1685 ----- 1685  
Db 80553 ACCTCTAGCGCTGAGGTCAAGCAAGGTCTATTTTAGGAATGTTGGCCCTTCAGAGAGT 80494  
QY 1685 ----- 1685  
Db 80493 TTTTCTTACTGAGCTATTCTGCCCCCTGTCTGTCTAGATTTTGGAACTTCTTCACATGG 80434  
QY 1686 -----GlyThrLeuGlnLysPheValAspAspLeuPheGluThr 1698  
Db 80433 TTCTTTTACCATCACTATAGGCGACGTGTCAGAAAGTTTGTAGATGACTTGTGTGAACCT 80374

QY 1699 IlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAsp 1718  
Db 80373 GTGTTTCAGTACAGCCACCGGGCTCAGCCTTACCTTGGCCATCAAGTACATGTTGAC 80314  
QY 1719 PheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrp 1738  
Db 80313 TTCTTGGATGAACAGGCTGACAGCGCAGATCAGTGCACCTGATGTGCGTCACACCTGG 80254  
QY 1739 LysSerAsn----- 1741  
Db 80253 AAGAGCAACTGGTACTTACTGGTCTGACTGCTGCTGGCTCTTCTGAGGGTGGGTGGGA 80194  
QY 1741 ----- 1741  
Db 80193 GGAAGAGCTAGTTACTTAGCCCAAGCCTCCAGGCCAATGGAGACAGGTTGACTAGCACT 80134  
QY 1741 ----- 1741  
Db 80133 GAATAGCCTTCCATGACAGAGGGATGGACAGTGTATGGCCTGGAAGTGAATATGAGTA 80074  
QY 1741 ----- 1741  
Db 80073 TATAGAAGTGGTCTGCTGGTGTGCGAGATGGGAAGTGTCTAGTCTAGCACTGGGAT 80014  
QY 1741 ----- 1741  
Db 80013 GACCATTCTTCTAGGTCTGTCAGGGTAGAGACACTGATGGCTGGCATAGGTGGA 79954  
QY 1741 ----- 1741  
Db 79953 GACAGTGGAGTTTCAGGGGTGGAGTTGAGGACAGAAAGCCATCTATTCTTATTCTCA 79894  
QY 1741 ----- 1741  
Db 79893 TTGCTGAGTGCAGCAGAGAGTTAAGAAAAACCCATTTATCCAAAATAGTCTCTCTGT 79834  
QY 1741 ----- 1741  
Db 79833 TACTTCAGGACAAGATATGACCTGCGCCAAACAGACAGGCTGTGAAGGAAGGCGTGGG 79774  
QY 1742 -----CysLeuPr 1744  
Db 79773 GAAGGGGCACATAGGGCCATTAGGAGTCTCAAGTGTATGCCCTCTCTTTCAGCTTACC 79714  
QY 1744 OLeuArgPheTrpValAsnMetIleLysAsnProGlnPheValPheAspIleHisLysAs 1764  
Db 79713 TCTGGTTCCTGGGTGAATGATGATCAAGATCCGCAATTTGGTTTGACATCATCAAGAA 79654  
QY 1764 nSerIleThrAspAlaCysLeuSerValValAlaGlnThrPheMetAspSerCysSerTh 1784  
Db 79653 TAGCATCACAGACGCTGTTTGTGTCAGTGGTGGCCAGACCTTTCATGACTCTCTGTCTAC 79594  
QY 1784 rSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysAspIl 1804  
Db 79593 ATCAGACACCGCTTGGCAGAGGACTACCTTCCAACAAGCTCTCTATGCCAAGATAT 79534  
QY 1804 eProSerTyrLysAsnTrpValGlu----- 1812  
Db 79533 TCCCAATTACAGAGCTGGGTGGAGAGGTGGGCGCTGGGATACTGCTGTGTATGCGCTGGT 79474  
QY 1812 ----- 1812  
Db 79473 GGCAGGAGGAACCTGCCATAACAGCCTGTCTGAGGCTGGCCAGATACAGAAAAAGCCCTC 79414  
QY 1812 ----- 1812  
Db 79413 AGAGCTGGCCAAACACTGATGGCATGAGTTTGGCATTTATGACCTTTCATGATTCAC 79354  
QY 1813 -----ArgTyrTyrSerAspIleGlyLysMetProAlaIleSerAsp 1827  
Db 79353 TCACCCAGGCTCTCTATAGGTACTATCGAGATTCGAAAGATGGCATCCATCAGTGACC 79294  
QY 1827 InAspMetAsnAlaTyrLeuAlaGluGlnSerArgMethHisMetAsnGluPheAsnThrM 1847



```

/rpt_family="CpG Island"
complement(11884..12163)
/evidence=experimental
/rpt_family="Alu-Sb0 or Alu-Sb1"
complement(12195..12482)
/evidence=experimental
/rpt_family="Alu-J or Alu-S"
complement(12668..12945)
/evidence=experimental
/rpt_family="Alu-J or Alu-S"
complement(18360..18648)
/evidence=experimental
/rpt_family="Alu-Sb0"
complement(18984..145006)
/gene="FLN"
complement(18984..19488)
/gene="FLN"
/number=48
/evidence=experimental
complement(join(19301..19488,19814..20017,20101..20319,
20483..20659,21361..21493,22033..22148,22336..22473,
22633..22899,23005..23127,23224..23376,23453..23656,
23748..23909,24006..24179,24367..24495,24603..24743,
24833..24935,25068..25163,25277..25524,27701..27724,
27884..28073,28649..28805,28895..29018,29434..29604,
29696..29856,29934..30096,30182..30355,30440..31037,
31758..32020,32120..32237,32429..32598,32692..32782,
32868..33028,33111..33234,34472..34615,34709..34822,
34976..35169,35271..35407,35586..35709,35798..35935,
36473..36673,36757..36919,37011..37088,37181..37299,
37846..37993,38090..38187,38291..38539,41319..41691))
/gene="FLN"
/codon_start=1
/evidence=experimental
/product="filamin"
/protein_id="AAA2644.1"
/db_xref="GI:1203969"
/translation="MSSSSRAGQSAAGAAPGGGVDRDAEMPATEKDLAEDAPWKI
QNTFTWCHNLKVCSSKRIANQTLDSGLRLIALLEVLSSQKMRKRNQRPTEFROM
OLENSVAFLEFDRESIKLVSDSKAIDVGNLKLGLIWLTLIHSYSPMPDDEED
EEAKQTPKORLLGWIQNKLPOLPTNFSRDMQSRALGALVDSACPLCDPDSWDA
SKPVTNAREAMQADHLLGIPQVITPEIIVDPNVDHSMVYILSQPKAKLKPAPLR
PLNPKRATAGCIEPTGNMVKRAEFTVETRSAGOGVELVYVVEDPAGHOEAKVTA
NDNKRFTSVMYPEVETGTHVTLFAGQHIKSPFVYVDKSGQDASKVTAQGPGL
PSGNIANKTTFIFFTAGTGEVEVVIQDPMGQGTVEPQLEARGDSTYRCYQPTM
EGVHTVHTTAGVPIPRSPYTVTVGOACNPSACRAGRGLOPKGRVRYKETADEFVYTK
GAGSGELKTVTKPKGEERVKOKDLGVYGEFYPMVPGTYIVTITWGGONIGRSP
EVVYVTECGNQKVRANGPGLGGVYVGSADFVEAIGDDVGTGLFSVEGPSQAKIECD
DKGGSCDVRIFQEAEGYAVHVLCSNEDIRLSPFMADIRDAPODFHPDRVKARGPGL
EKTGVAVNKPAAFTVDKAGHGKAPLRVQVQDNEGPVEALVKONGNTYSCSYVPKPK
VKHTAMVSGVSPINSPFRVNVGAGSHPNKVKYVGPVAKTGLKHAEPYTFVDCAE
AGQDVSIGIKCAPGVVPAEADIDFDIIRNDNTFTVKYTPRAGAGSTYIMVLADQAA
TPTSPIRVKVPESHDSKVAEGPLSRTGVELGKPTFTFTNAKAGKGLDVQFSL
TKGDVARDVLDLIDHNDNTYVKTPTVQOQPGVNVNVTGGDIPKSPSPVAVSPSLDLS
KIVYGLGERGVKQDEFTVKSAGGQGVKASKIVGPSGAAPVCPKVEPGLGADNSV
VRLEPREGEYEVETVVPVPGSPFPFLEAVTPKPSKFAFGPQLOGSGAPARF
TIDTKAGTGLGLTVGPGCEAQLCDLNGDGTCSVSVPTPEPDYINILFADTHIP
GSPFKAHVWPCFASKVKCSGGLERATAGEVGQFQVDCSAGSAELTIEICSEAGLIP
AEVYIDHGDGTHITITVILPCPGATVTVIKYGOPVFNPSKLOVEPAVDTSVQCYC
PGLGOGCFREATEFSDARALTQTGCPHVHKAIVANPSGNLTITYYVQDRDGMKVE
YTPYEGLSHVDVITDGSPPFPVPTGDCPSRVRVHGPGLQSGTNNKPKTV
ETKAGTGGGLGAVGSEAKMSMDKNDGSCVEYIPYEAGTYSLNVTYGGHVPGS
PFKVPVHDVTDASKVKCSGPGLSGPMVRANLPQSFDVTSKAGVAPLQVKVQGPGLV

```

## Alignment Scores:

Pred. No.:	0	Length:	219447
Score:	4394.00	Matches:	1127
Best Similarity:	36.84%	Conservative:	297
Best Local Similarity:	29.16%	Mismatches:	423
Query Match:	43.98%	Indels:	2023
DB:	9	Gaps:	47

US-09-964-956-13 (1-1896) x HUMFLN66PD (1-219447)

```
QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrylsGlu 273
|||||
Db 131636 GTGACTTCTGAGCTGACGACACCCAGCAGAGCTGTGGACACAGCGGCGAG 131695
QY 274 GlnValTyrThrSerLysLeuValArgLysCysLysGluAspThrAlaPheAsnSerTyr 293
::: |||||
Db 131696 AAATTTTTCAGCTCAAGATGCTGCGATGTCGCGGAGACTCAGAGATTCTACTCATAC 131755
QY 294 ValGlnValProIleGlyCysGluArgSerGlyValGluTyrArgLeuGlnAlaAla 313
|||||
Db 131756 GTGGAATTCCTCCATCGCTGCTGCGCGCGCTGGAGTACCGCTTGGTCAGAGCGCC 131815
QY 314 TyrLeuSerLysAlaGlyAlaValLeuGlyAlaValLeuGlyValHisProAspAspAsp 333
::: |||||
Db 131816 CACCTGGCAAGCTGGCTGCTGCTGCGCCAGCGCTGGCGCTGGCGCTGATGAGGAC 131875
QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLysAspGluSer 353
::: |||||
Db 131876 GTCCCTCTTCACCATCTTCTCTCAGGGCCAGAGAACCGGGCCAGCCACCGCGCAGACC 131935
QY 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
|||||
Db 131936 ATCCTCTGCTCTTCACCTCAGCAACATCAATGCCACATCGCGCGCATCCAGTCC 131995
QY 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
|||||
Db 131996 TGCTATCGTGGGAGGCACTCTGGCTCTGCCCTGGCTGCTGAACAAGGAGCTGCCCTGC 132055
QY 393 ----- 393
Db 132056 ATCAACACCGTGAACCCCTCATCACCCACACTGGTCTCTGCCCTGTCCAGGTCTACC 132115
QY 393 ----- 393
Db 132116 ATGCCACCGTGGGCTCAGCGCCAGTCTATCTGTCCAGGCTTGCCATGGCCCTGGGAG 132175
QY 393 ----- 393
Db 132176 TGGCACCTTTGCCCTCCACCTGTTCTCTCCCTCCAGGGTCCCTGCCCTCTCCCTGGCT 132235
QY 394 -----SerSerAlaLeuLeu 398
|||||
Db 132236 CTCTCTGGGCTGCCAGTGGCCAGCTTCTGCTTCTGCTTCCCTGCTCCAGCTGTTCTCTCT 132295
QY 399 Thr ----- 399
Db 132296 GCTCCGTCTCCCGTTGGCCCTTGTCTTGTAGCCAAAGAGTCTTGTGGGCCAGGCTTC 132355
QY 399 ----- 399
Db 132356 GCTCCTCGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132415
QY 399 ----- 399
Db 132416 ACCTAGGGCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132475
QY 399 ----- 399
Db 132476 CTGGGACTCAGGAACCTGGCGCTCCATCTTGGCGCTGGGAGCTTGTACTCTCAGGGTT 132535
QY 400 -----IleAspAsnPheCysGlyLeuAspMetAsn 410
|||||
Db 132536 CCTCCTCTGTTTACCAGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 132595
QY 411 AlaProLeuGlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAsp 430
|||||
Db 132596 CAGCCTCTGGAGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 132655
QY 431 ArgMetThrSerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThr 450
|||||
Db 132656 GGCATGGCAGCGTGGCGGCTTACACCTACCGCCAGCAGTCTGTGTGTCTTCTATGGCAGC 132715
QY 451 LysSerGlyLysLeuLysLys ----- 457
::: |||||
```

```
Db 133796 CATCAGGTGGTCTGTGGCTCAGGTGGGGTGGATGGCTTCAG---GATGCCACCTG 133852
QY 470 TyrGluThrValGlnValValAspProGlyProValLeuArgAspMetAlaPheSerLys 489
Db 133853 TATGACAGTCCCGTGTGGATGCGACGCCCATCTCCGAGACCTGCTTTCAGCCCG 133912
QY 490 AspHisGluGlnLeuTyrIleMetSerGluArg----- 500
Db 133913 GACCACCGGCACATCTATCTCTGAGTGAGAACAGAGTGGCGCTGTGGTGGGGTG 133972
QY 501 -----GlnLeuThrArgValPr 506
Db 133973 GCTGTGGGGGGGGTGGGTGGGTGGTGTGGCTGTCCTCCAGGTGAGCGAGCTCCC 134032
QY 506 oValGluSerCysGlyGlnTyrGlnSerCysGlyGluCysLeuSylSerGlyAspProHi 526
Db 134033 GGTGGAGACCTGTGAGCAGTACAGAGCTGCGAGCGTGCCTGGGCTCGGGGACCCGCA 134092
QY 526 sCysGlyThrCysValLeuHisAsnThr----- 535
Db 134093 CTGTGGTGTGTGCTGCTCGCACAG -GTGAGGGCGGGGCCCTGCTCGGGGAGTTGG 134151
QY 535 ----- 535
Db 134152 AGGGCCCCACTGGCCAGGGGGAGCCAGCCACACCCAGCGCTGGCTTGGGGCTCCCCCC 134211
QY 535 ----- 535
Db 134212 GCCCTCCTCCACTTCTCCAGCTAATGAGTGGNAACCTCCACCCCGAGTGACGTCCCTCGG 134271
QY 535 ----- 535
Db 134272 GCCCCAGGGGACGGGCCAAGGCTGCTGTGCTGGCACGTGAGCTGGACAGGCGCTGG 134331
QY 536 -----CysThrArgLy 539
Db 134332 GCCCTCCCGGGGAGTGGGGGACCGGCTCTGGCCCGACCCCGTGCAGGTGCTGCCGCGA 134391
QY 539 sGluArgCysGluArgSerLysGluProArgPheAlaSerGluMetLysGlnCysVa 559
Db 134392 AGGGGCTGTCTGGGGCTGCTGCCCCACACAGCTTTGTGAGGAGCTGAGCAAGTGTGT 134451
QY 559 lArgLeuThrValHisProAsnIleSerValSerGlnTyrAsnValLeu----- 577
Db 134452 CCAGGTGGGGGTCCGGGCCCAACAATGTGCTGACGTACCTTGGGGTGCAGGTGAGCAG 134511
QY 578 -----ValLeu-GluThr----- 581
Db 134512 CTTGGGGTGGCGGCTGGGTGTGCATGTGTGCTGGGAGTCCGCCCTGCCCTGAGCCC 134571
QY 582 -----TyrAsnValProGluLeuSerAlaGlyValA 592
Db 134572 TCTGTCTCCCGCAGCTGACCGTCAACCTGCACAAGTGCAGACCTCAGTGGCGGTGA 134631
QY 592 sNCysThrPheGluAspLeuSerGluMetAspGly---LeuValValGlyAsnGlnIleG 611
Db 134632 GCTGGCCCTTCAGGCGGGCGGAGAGACGAGCGGTCTGCTGCCCTCCGGTGAACCTGC 134691
QY 611 lncCysThrSerProAlaLaLysGluValProArgIleIleThrGluAsnGlyAsp--- 629
Db 134692 TCTGCCCTCAACCTCCCTCCAGGAGCTCCGAGCTCTTACCAGGGGCGATGGTCAGTGGG 134751
QY 629 ----- 629
Db 134752 TTGGGGCTGCCAGGATGGGCAGAGTGGGGCTCTCCCTACCCCGCAGGATTCACGGC 134811
QY 630 -----His-ValValGlnLeuGlnLeuLysSerLysGlu 641
Db 134812 CACCCCGGAGCTGCTGACGGGGCCACCCGCACTGTGCGGTGTCAGCTTCTCTCAAGGAG 134871
QY 642 ThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerValHisAsnSer 661
Db 134872 ACAGGCGTGGGTTGCGCGTGTGCTTGTCTTACAACTGCAGCGTCTCCAGTGC 134931
```

```
QY 661 ----- 661
Db 134932 TGAGTACTGGCAGCACCCGCTCCCTACCCCTGGGATAGAGGGACCTCTCGGACAGCC 134991
QY 661 ----- 661
Db 134992 GACACCTTACGGGAGCTGTTTCTGGGATCAGGGGCTAACTGTACGGCCCTGATAGCC 135051
QY 661 ----- 661
Db 135052 CCTGACATCCCATCATCTCTCTGCTCCCTGATATGGCTGCCTTCTTCAGTTCCCGA 135111
QY 661 ----- 661
Db 135112 TGTGCTCTATTGGGAGCAGTGAGGGGAGTGCAGCCCTCTCTGTTATTCTTGAAGCCAC 135171
QY 662 -----CysLeuSerCysValGluSerProTyrArgCysHisThrCysLysTyr 677
Db 135172 TTCCCCCCCCAGGTGCATGCTCTGTGTTGGCAGCCCTTACCCCTGCCACTGTGTAAGTAC 135231
QY 678 ArgHisValCysThrHisAspProLysThrCysSerPheGlnGluGlyArgValLysLeu 697
Db 135232 CGCCACAGTGTACAGCGGCCCCACAGTGTCTCTTCCAGGAGGCGAGGTCCACAGC 135291
QY 698 ProGluAsp----- 700
Db 135292 CCTGAGGT -GAGGCGGGGCGCCATGTGAGGGGCTGGGCTCTGTGTCGGGCGGGGCCA 135350
QY 700 ----- 700
Db 135351 CCGGCTTCTATGCGTCTCTCGGTTTCTTTTGTAGCCTTCTCCAGTTGGGCTGGAGAGGT 135410
QY 701 -----CysProGlnLeuLeuArgVal 707
Db 135411 GGCCCTGTGCAGCTGACAGTGTGTTTCCCGCAGAGGCTGCCCTGAGATCTGCCACGT 135470
QY 708 AspLysIleLeuValProValGluValIleLysProIleThrLeuLysAlaLysAsnLeu 727
Db 135471 GGGGACCTCTGTATCCCGGTTGGGTGCATGAGCCCTCTTACCTCGGGCTAAGAACCTA 135530
QY 728 ProGlnProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGlu 747
Db 135531 CCTCAGCGCAGTCCGGCCAGAACATATGAGTGGTGGTGGGTCGAGGGGGGGCAG 135590
QY 748 GlnArgValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThrSer--- 766
Db 135591 CAGCGGTGCTCGCGTGGCTTCAACAGCAGAGTGTGACGTGCCAGAACCCCTCGGTG 135650
QY 766 ----- 766
Db 135651 AGGTCCACCGCTGCCCTCCCTCGGGCTGCGGCTGTGGCTGTGGATCGTTCTTG 135710
QY 766 ----- 766
Db 135711 TCTATGCCAGCTCTCTGGCAGGAACTGTCTGAGTCTCTGCTAGGATTCCTGTACCT 135770
QY 766 ----- 766
Db 135771 GGAGGAGGGGCGCAGCTGGCAAAATCTTGGTAGGGGTTCTGCACATCTTATCTGGGGT 135830
QY 767 -----TyrSerTyrGluGlyMetGluIleAsnAsnLeuProVa 779
Db 135831 CCATGGGTCTCTTCTCCAGTACTCTCTATGAGGTGATGAGCATGTGTGACACCCAGCT 135890
QY 779 lGluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsnProAlaGln-AsnLysV 799
Db 135891 GGAGTTCTCGGTGTGGGTGAGACTTCCCATAGACAAGCCTCCAGCTTCCGAGT 135950
QY 799 alHis----- 800
Db 135951 TGAAGGCATGGGCCGAGGAGCTTCCTCTTAAGGCAATTGGCACTGCTGGGGTCTCGGGTCT 136010
```

```
QY 800 ----- 800
Db 136011 GGGGGCCCTAGTGCTGCCACCTTTTCCACCTTTTTCGCCACCTGCTGCTCCGTCAGCAGTGC 136070
QY 801 -----LeuTyrLysCysGlyAlaMetArgGluSerCysGlyLeuC 814
Db 136071 CTTCTGTGCTGTCAGCCCTCTGTACAAAGTCTGGGCGCAGCGCCACCTGTGGCCCTCT 136130
QY 814 ySLeuLysAlaAspProAspPheAlaCysGlyTyrPcysGlnGlyProGlyGlnCysThrL 834
Db 136131 GCCTCAAGGTGATCCCGCTTCAACTGTGGTGTGATCTCAGAGCACAGTGCACAGC 136190
QY 834 euArgGlnHisCysProAlaGlnGluSerGlnTrpLeuGluLeuSerGlyAlaLysSerL 854
Db 136191 TGGGACCACTGCGCGCCCGCAAGAACCACTGGATGCACCTGAGCCAGTGCAGAGGCGCACCC 136250
QY 854 yScysThrAsnProArgIleThr----- 861
Db 136251 GGTGCAGCCACCCCGCATCACAGGATCAGCCTCCCTTCACGGCCCTGCGCCACTGCGCAA 136310
QY 862 -----GluIleIleP 865
Db 136311 CAGGGCCCTGGAGTCTGAGCCAACTCTCTCACTGCCATCTCTCCACAGATCCACC 136370
QY 865 roValThrGlyProArgGluGlyThrLysValThrIleArgGlyGluAsnLeuGlyL 885
Db 136371 CTCTCGTGGGCCCCAAGGAGGAGGACCCCGGTCACTGCTGGGTGACAACTGGGCC 136430
QY 885 euGluPheArgAspIleAlaSerHisValLysValAlaGlyValGluCysSerProLeuV 905
Db 136431 TCTTGTCCGAGAGGTGGGC-----CTGCGGGTGGCTGGCGTGTGCAACTCCATTC 136484
QY 905 alAspGlyTyrIleProAlaGlu----- 912
Db 136485 CGGCGGAGTACATCAGTCTCAGAGGTGAGTGGCGCTCTGTGGGTGCGCGGCGCTATGT 136544
QY 913 -----GlnIleValCysGluMetGlyGluAla- 921
Db 136545 GGCCTGCGCGCCCTGACGCTCTCTGAGCCCTAGGATCGTGTGTGAGTGGAGAGTGC 136604
QY 922 --LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgProGluP 941
Db 136605 TGGTGGCCAGCCCGCGCGGCGCTGGAGCTGTGTGGTGACTGTTCAGCCGACT 136664
QY 941 heMetAlaArgSerSerGlnLeuTyrTyrPheMet----- 952
Db 136665 TCCGACGACGTGCGAGCAGCTACAGCTTTGTGGTGGTGGCTGCGCGGCGCTACCCCT 136724
QY 952 ----- 952
Db 136725 TCCTGTCCCTCTCTCTCCGCAAGGGCGGTGTGGAGAGCGCGCGCGCTCTCCCT 136784
QY 953 -----ThrLeuThrLeuSerAspLeuLysProSerArgGlyProM 966
Db 136785 CAGGGCAGCTTCTCCCGCAGACCCCAACGTTTGACCAAGTGAAGTCCCGAGCGGTGCGCGG 136844
QY 966 etSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnValV 986
Db 136845 CGTCCGGGGCACACGGCTTACCATCTCAGGAGCTCTGTGATGCTGCGCAGAGGTCA 136904
QY 986 alValMetPheGlyLysGlnProCysLeuPheHisArg----- 998
Db 136905 CAGTGACTGTGAGGGACAGCAGTGCAGTTTGTAAAGTGTGGCGCGCGGCGCTGCCAGCTT 136964
QY 999 -----ArgS 1000
Db 136965 TGGGTGGGCATCTGTGGGGCGCTGGGACGGGTGGCTGAGGGCCCTGGGCCACCCG 137024
QY 1000 erProSerTyrIleValCysAsnThrThrSerSerAspGluValLeu----- 1015
Db 137025 CTCAAGCAGCCCTGCT--TGCCAAATGAGGAGATGCCAAGGCGATCGTGTGATCTCTCAC 137083
QY 1016 -----GluMetLysValSerValGlnValAspArgAlaLysI 1028
```

---

```
Db 137084 CTCTCTCCACCTGGGCCCCACCGAGGCCCATCACATCTGCGATTCGCGGCTAAACA 137143
QY 1028 leHisGln--AspLeuValPheGlnTyrValIleuAspProThrIleValArgIleGluP 1047
Db 137144 TCTCCAGCCCGGGTCTATCTACACTACTCAGAGCCCGCCACCGCTTGAGC 137203
QY 1047 roGluTrpSerIleVal----- 1052
Db 137204 CCACCTGGAGCATCATCAAGTAAGACCTTGGGGAGCTGGGAGCCTGGCAGTGTCCAGAG 137263
QY 1052 ----- 1052
Db 137264 GGCCTCAGGAGTGGGTGCTCTGAGCTCCGFTGGGCCCCCTCTCTGAGCCTAGGCGCTC 137323
QY 1052 ----- 1052
Db 137324 ATTGGTGGAGGGTGGAGCTGTCTGTGGTGCACAGTACCCGGCACCCACTAGGCGATCC 137383
QY 1052 ----- 1052
Db 137384 AGGGTCAGGGAAGGCTGGCTGCATGGCAGCCTTGATCGCTCTCCAGGCTGGTGGT 137443
QY 1053 -----SerGlyAsnThrProIleAla 1059
Db 137444 GCACCTGGAGAGGAGCGCTGAGGCCCTGCTGCTGCTCCCTAGTGAAGACACTGCCATCCT 137503
QY 1060 ValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGlyGly 1079
Db 137504 GTGAGTGGGACCCACTGCTGACGGTCCAGGAGCCCGGGTCCGTGCCAAGTACCGCGGC 137563
QY 1080 LysGluHisIleAsnIle----- 1085
Db 137564 ATTGAGACCACCAATGTGATACAGCTGCCCGCCCCACCCGACCCCTGCAGCCCATGG 137623
QY 1086 -----CysGluValLeuAsnAlaThr 1092
Db 137624 CTTACGCTGCTGGTGTCCACCTTTGTCCCCACACAGTGCCTCAAGTATCAAGACACT 137683
QY 1093 GluMetThrCysGlnAlaProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThr 1112
Db 137684 GCCATGCTGTGAAGCCCGCCGCACTTTCTTGGCGGCGCCCGCCAGCTCGGCGCAAGGC 137743
QY 1113 GluArgProGluGluPheGlyPheIleLeuAsnValGlnSerLeuLeuIleLeuAsn 1132
Db 137744 GAGCACCTGTATGAGTTGGCTTCTGCTGGACACAGTGCACAAACGGCCCGCTCCCTCAAC 137803
QY 1133 LysThrAsnPheThrTyrProAsnProValPheGluAlaPheGlyProSerGlyIle 1152
Db 137804 CGCTCTCTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 137863
QY 1153 LeuGluLeuLysProGlyThrProIleIleLeuLys----- 1164
Db 137864 CTGGAGCTCAACCGGGCTCCACGTTGCTGAAGTGTGGCGGGGTGGGGCGGGGA 137923
QY 1165 -----GlyLysAsnLeu 1168
Db 137924 GGGCGGGAAGTGGAGAGTCTCTGGGCTGAAGTTGCTCTCCACCCCGGAGCAAGACCT 137983
QY 1168 uIleProValAlaGlyGlyAsnValLysLeuAsnTyrThrValLeuValGlyGluLys 1188
Db 137984 GATTCCCGCGGACCGCGGAGCTCC--CGCTCACTACTACTGTGTGTATAGGAGGCGCA 138040
QY 1188 sProCysThrValThrValSerAspValGlnLeuLeuCysGluSerProAsnLeuIleG 1208
Db 138041 GCGGTGCTGCTCCTGCTCGGACACACAACTCTCTGCGAGCTCACTCCAGCCAGAGCTGG 138100
QY 1208 yArgHisLysValMet----- 1213
Db 138101 CCGGACAGCTGTCTATGATAGGTGGGATGGGAGACCCCTTGGCGACCCCGAGGTGGCG 138160
QY 1214 -----AlaArgValGlyGlyMetGluTyrSe 1222
```

Db 138161 TGGTGGTACGCTACCTCAGCGCTGTCTCCACAGGTGCTGGTGGCTGGAGTCTGTG 138220  
QY 1222 rProGlyMetValTyrIleAlaProAspSerProLeuSerLeuProAlaIleValSerII 1242  
Db 138221 GCTGGGCACCTCGACATCTCGCAGAGCGGCGCTGACCTACCGCCATGATGGGGCT 138280  
QY 1242 eAlaValAlaGlyCysLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrLysAr 1262  
Db 138281 GCGCGGGGGGTGGGCTCTGTGTGGCATACACAGCGGTGTGTGGCGTACAAGCG 138340  
QY 1262 gLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuG1 1282  
Db 138341 CAAGACTCAGACGCGGACCGTACCTCAAGCGTCTCGAGTGCAGATGGACAACCTGGA 138400  
QY 1282 uSerArgValAlaLeuGluCysLysGluAla 1292  
Db 138401 GTCCCGTGTGGCTGGAGTGAAGGAAGG-TGCGTTCAGGCGGGGGGATGTGGTGG 138459  
QY 1293 1293 -----PheAlaG1 1295  
Db 138460 AAGCTGGGGACCTCCCTCTCCACACTCATTCCTCTCCACCCCGCAGCTTTTGCAGA 138519  
QY 1295 uLeuGlnThrAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLe 1315  
Db 138520 GCTGCAGACGGACATCAATGAGCTGACTAACCATGGACGAGGTGCAGATFCCCTTCCT 138579  
QY 1315 uAspTyrArgThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLe 1335  
Db 138580 GGACTACCGGACTTACCGCTGCGGTGCTCTTCCCGGGCATCGAGGCCACCGGTGCT 138639  
QY 1335 uArgAspLeuGluVal-----ProGlyTyr 1343  
Db 138640 CAAGGAGCTGGATGTGAGCTCTGCTGGCTGCGCCCGCCACCATTCCTCTCAGGGCGGCC 138699  
QY 1344 -----ArgGlnGluArgValGluLysGlyLeuLy 1353  
Db 138700 CCACACCTCTGAGACCTCTGCTCCCGCACAGACCCCAACGTGGAGAGGCCCTGCG 138759  
QY 1353 sLeuPheAlaGlnLeuIleAsnLysValPheLeuLeuSerPheIleArgThrLeuG1 1373  
Db 138760 CCTCTTGGGCGCTGTCACAGCGCGGCTGCTTACCTTCATCCACACGCTGGA 138819  
QY 1373 uSerGlnArgSerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrVa 1393  
Db 138820 GGCCCCAGAGAGCTTCTCCATGCGCGACCGCGCACCGTGGCTCGCTCCACATFGGTGGC 138879  
QY 1393 lLeuGlnSerLysLeuGluTyrAlaThrAspValLeuLysGlnLeuAlaAspLeuI1 1413  
Db 138880 CTTGCAGAGCGGCTGACATATGCCAGGGGCTGCTCAAGCAACTGCTGGCCGACCTAT 138939  
QY 1413 eAspLysAsnLeuGluSerLysAsnHisProLysLeuLeuArg----- 1428  
Db 138940 CGAAGAAGCTCGAGAGCAAGAACCCCAAGCTGCTGCTAGCGAGGTACCTGCCTCG 138999  
QY 1428 ----- 1428  
Db 139000 CTCTATCCAGGCCACACCTTTGTCTGCTGGGCGCTGCCCGCTGCCAAGCCCGCCCT 139059  
QY 1429 -----ArgThrGluSerValAlaGluLysMetLeuT 1439  
Db 139060 GCCAGGCCCGAGCCCTCTCTCTGCTGCTAGGACAGAGTCAAGTGGCTGAGAGATGCTTA 139119  
QY 1439 hrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLys----- 1451  
Db 139120 CCAACTGGTTCACGTTCTCTGCTGATAAGTTTCTGAAGGTGCGCGCAGGTGGTGGCGGC 139179  
QY 1451 ----- 1451  
Db 139180 AGGAGGTGTGGCAGAGGACCGGCCAGTGTGTGAGCGAGGAGGTGCGCCAGCGAGGC 140315  
QY 1452 -----GluCysAlaGlyGluProLeuPheSerLeuPheSerAlaIleLysGlnG 1468  
Db 139240 CCTGGCTGACGAGAGTGTGCTGGGAGCGCTCTCTCTCTGCTTACTGTGCCATCAAGCAGC 139299

QY 1468 lMetGluLysGlyProIleAspAlaIleThrGlyAlaArgTyrSerLeuSerGluA 1488  
Db 139300 AGATGAGAGAGGGCCCGCCATGATGCCATCACGGCGAGCAGATACTCCCTCAGCGAGG 139359  
QY 1488 sPLeuLeuIleArgGlnGlnIleAspTyrLysThrLeuVal----- 1501  
Db 139360 ACAAGCTCATCGTCAGCAGATCGACTACAAGACACTGGTGAGCGCAGGGCCAGCGGGC 139419  
QY 1501 ----- 1501  
Db 139420 CAGAGGTAGGGCGCGCAGAGAGAGGCTCGCCCCAGACTGACACTGGAGTCCGCTTTCGCC 139479  
QY 1502 -----LeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLysI 1519  
Db 139480 TCAGACCCCTTTCAC-TGCGTGTCTCGGAGACAGGCGCAGCGCCAGGTCCCACTGAAG 139538  
QY 1519 lLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLysA 1539  
Db 139539 TTCTCAACTGTGACAGCATCACCCAGGCCAAAGATAAGCTGTGGACACTGTGTACAAGG 139598  
QY 1539 snValProCysSerHisArgProLysAlaAlaAspMetAspLeu----- 1553  
Db 139599 GCATTCCGTTACTCCAGCGTCCCAAGCTGAGACATGAGCCTGGGTGAGGTCCCCACCC 139658  
QY 1553 ----- 1553  
Db 139659 TCTCCTCTGGCTCCCACTCATACCTCTGTCGCCGTGTGACCTCCGCTGGCTCTCCCGC 139718  
QY 1553 ----- 1553  
Db 139719 TCCCCACCTGAACCTCTGTTCTTAAGAGAGAGTGGCGCAGCACCTGGGGGCTGCCAGCAT 139778  
QY 1554 -----GluTipArgGlnGlySerGlyAlaArgMetIle 1564  
Db 139779 AATCTTAAGGCTGTCTGTGGCCCCACAGAGTGGCGCCAGGCGCGCATGACTCCATCATC 139838  
QY 1565 LeuGlnAspGluAspIleThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeu 1584  
Db 139839 CTCCAGGATGAGGATGTCAACCAAGATCGAGTGTGACTGGAAGAGGCTCAACTCAGCTG 139898  
QY 1585 AlaHisTyr----- 1587  
Db 139899 GCGCCACTACAGGTGAGGGTGGGGGCCCATTCCTCCCGAGGGCCACCTGGGAGCCAGGA 139958  
QY 1587 ----- 1587  
Db 139959 CCCTGCCTTGAGCTCCAGCAGACAGCGGAGCAGTGGCCCTGGCTCCCTCGCCCTTCTC 140078  
Db 140019 CTGGGCTCGTGGGCTCCCTCTCTGGGTGGGTGGGGCGCTTGGCTTCTCAGC 140078  
QY 1588 -----GlnValProAspGlySerValValAl 1596  
Db 140079 TTCTCGACTGCCCCCTCTGTCTGGGCGCTCCAGGTGACAGCGGTTCCTTGTGGC 140138  
QY 1596 aLeuValSerLysGlnValThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgTh 1616  
Db 140139 ATTGGTGCCCAACAAGTGTCTGCCTATPAACATGGCCAACCTCTTCACTTCCACCGC 140196  
QY 1616 rSerAlaSerLysTyr----- 1621  
Db 140197 -TCCCTCAGCCGCTACGGTAGGTGTCTCAGTGTGTGGTGGCCATGTGCCCTTTCAGGGAAC 140255  
QY 1621 ----- 1621  
Db 140256 CCCCCTTCCAAAGTGCATCGATTCTGTAGATGTAGACGGAGGTGCGCCAGCGAGGC 140315  
QY 1621 ----- 1621  
Db 140316 AGATGGGTCCCCACATGCTGTGAGCTCCCGGGAGAGTGGGCGAGGGGCCAGGTGTGGC 140375

```
QY 1622 -----GluAsnMetIleArgTyrThrGlySerProAs 1632
Db 140376 CTAGGGTCACATCATCTCTGCTCCAGAGAGCTTGTCTCGCAGCGGCAGCCCTGA 140435
QY 1632 pSerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyValLysMetTr 1652
Db 140436 TAGCCTCGGCTCAGCGGCACCATGATATACGCTGACCCAGGAGACAGCAGCAATTTGT 140495
QY 1652 pHisLeuValLysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGlySerLysMe 1672
Db 140496 GCACCTGGTGAAGAACACAGCACCATTGCGGCATTCGCGAGGGGACCGTGGCAGCAAGT 140555
QY 1672 tValSerGluIleTyrLeuThrArgLeuLeuAlaThrLys----- 1685
Db 140556 GGTCTCGAGATCTACCTGACACGCTGCTGGCCACCAAGTATATGGGCTGCTCTCGCC 140615
QY 1685 ----- 1685
Db 140616 ACCTTGGCTCCGCCAGAGGTTTCCCGGCTTACAGGGAGGGGCGCATGGATCTTG 140675
QY 1685 ----- 1685
Db 140676 CTTCCAGTGGCCCTTCTTCGGGCTTTGCTGTGGAACTCTGACCCCTAGATAAGGCC 140735
QY 1685 ----- 1685
Db 140736 ATGCTGTCTGAGAGACCACTGGCCCTGTAGGGAAGCCCTCTCTTTGCCAAGCCCTTCT 140795
QY 1685 ----- 1685
Db 140796 GCCTCCATCTCTCTGCTCTGGGACATCCCGACCTGGCTCCCTCATAGCCTGTGACCTC 140855
QY 1686 -----GlyThrLeuGlnLysPheValAspLeuPheGluThrIlePheS 1701
Db 140856 TCTGCCACACAGGACACTGCAGAAAGTTCGTGATGACCTCTTGAGACAGTGTCA 140915
QY 1701 erThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuA 1721
Db 140916 GCACAGCCACCGGGGCTCGCCCTGCGCCCTGCGCATCAAGTACATGTCGACTTCCTGG 140975
QY 1721 spGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrpLysSerA 1741
Db 140976 ATGAGCAGCGGACCGCGCCAGATCAGCGACCCCGATGTGCGCCACACCTGGAGAGCA 141035
QY 1741 sn----- 1741
Db 141036 ACTGTTATCACCCCGTCTGGCTGCCAGCAGCCTGTCTGGAGACTGTGTGGCGGAAGAC 141095
QY 1741 ----- 1741
Db 141096 CTTGGTGGCTGTGTGAGACCCAGGCGCTGGAGTAGATCCAGGAGGGAATGAGCC 141155
QY 1741 ----- 1741
Db 141156 CTGGCTGGGTTGACATGACGTTGCCATCAAGCAAGGGGCGCTTTCGTCGATCTCT 141215
QY 1741 ----- 1741
Db 141216 GGAGCTGAGCTCCCTGGGTGGGTGGATGTGTGTGTGAGGCTTTGCCCTCTGGGGTCT 141275
QY 1741 ----- 1741
Db 141276 TGCACCAAGTAGAGATGTGTGTGGTTCGCGGTGAGAGCGGTGAGAGGAGCAAGC 141335
QY 1741 ----- 1741
Db 141336 TCGCGAGGGGTTGAGCCCTATTGAGACAGAAGTTACGACAGACTTAGAGGAAGGCC 141395
QY 1741 ----- 1741
Db 141396 GTTACCCTAAATGTGTGTCTCTCACCACTCTGACCAAGATGTGAGCGCGCGGACAGG 141455
QY 1741 ----- 1741
```

---

```
Db 141456 AAAGAGATGTGCACATGGAGGGGGCCCTTGGGGCTAACTAGGAGCCTCAGGCGCAGCT 141515
QY 1742 -----CysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGln 1756
Db 141516 CCCTCTGTGTGTCCACAGCTGCGCGCTTCTGGGTGAATGTGATCAGAAGCCGACG 141575
QY 1757 PheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValAlaIaGln 1776
Db 141576 TTCGTGTTCACATCCACAAGAACACATCACGATGCTGCTGTCGGTGGTAGCCAG 141635
QY 1777 ThrPheMetAspSerCysSerThrGluHisArgLeuGlyLysAspSerProSerAsn 1796
Db 141636 ACCTTCATGAGCTCTCTCTACATCCGAGCACCGCTGGGGAAGGACTCGCCCTCCAAC 141695
QY 1797 LysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGlu----- 1812
Db 141696 AAATGCTCTACGCCAAGACATCCCACTACAGAGCTGGGTGGAGAGGTGGGCTCCG 141755
QY 1812 ----- 1812
Db 141756 CCCGCTGTGGTGGCAGAGGCGAGGACCTGCCCTCCCTGGGCTGCTGTGCGAGGACG 141815
QY 1812 ----- 1812
Db 141816 CCAGCTACAGGAGGAGGAGCCCGGGGCTGGCTGGGCGAGTGGGAGGCGGGTCTAGTTT 141875
QY 1813 -----ArgTyrTyrSerAspIleGlyLy 1820
Db 141876 TGTGATGTCGCTGAGTACATCTCAGGCTGCCACAGCTATTATCGAGACATTCGCAA 141935
QY 1820 sMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluGlnSerArgMetHi 1840
Db 141936 TGTGATGTCATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 141995
QY 1840 sMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGlyLysTy 1860
Db 141996 CGCCAGCGACTTCAGGCTGCTGAGTGGCTCAGGAGCTGATTTCTATGTCACCAAGTA 142055
QY 1860 rSerGluGluIle-----LeuGlyPro 1867
Db 142056 CGCCAGGAGGCTGTGTCTATCCACAGACTCCACGTTACTTGGTCT 142104

RESULT 14
AL592072/c
LOCUS
DEFINITION
Homo sapiens chromosome X clone RP13-156P1, *** SEQUENCING IN
PROGRESS ***, 12 unordered pieces.
ACCESSION AL592072
VERSION AL592072.2 GI:14530021
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Heath, P.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
COMMENT On Jun 22, 2001 this sequence version replaced gi:14456388.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: b8156p1
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator ET-amersham; 26% of reads Chemistry:
```

Dye-terminator Big Dye; 73% of reads  
Consensus quality: 156400 bases at least Q40  
Consensus quality: 158019 bases at least Q30  
Consensus quality: 158870 bases at least Q20  
Insert size: 159556; sum-of-ontigs  
Insert size: 142754; agarose-fp  
Quality coverage: 6.25x in Q20 bases; sum-of-ontigs Quality  
coverage: 7.30x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 17742: contig of 17742 bp in length  
\* 17743 17842: gap of 100 bp  
\* 17843 25453: contig of 7611 bp in length  
\* 25454 25553: gap of 100 bp  
\* 25554 29396: contig of 3843 bp in length  
\* 29397 29496: gap of 100 bp  
\* 29497 36990: contig of 7494 bp in length  
\* 36991 37090: gap of 100 bp  
\* 37091 57180: contig of 20090 bp in length  
\* 57181 57280: gap of 100 bp  
\* 57281 70963: contig of 13683 bp in length  
\* 70964 71063: gap of 100 bp  
\* 71064 78142: contig of 7079 bp in length  
\* 78143 78242: gap of 100 bp  
\* 78243 84646: contig of 6404 bp in length  
\* 84647 84746: gap of 100 bp  
\* 84747 104953: contig of 20207 bp in length  
\* 104954 105053: gap of 100 bp  
\* 105054 129704: contig of 24651 bp in length  
\* 129705 129804: gap of 100 bp  
\* 129805 151925: contig of 22121 bp in length  
\* 151926 152025: gap of 100 bp  
\* 152026 160656: contig of 8631 bp in length.  
Location/Qualifiers  
1..160656  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/clone="RP13-156P1"  
/clone\_lib="RPC1-13.1"  
1..17742  
/note="assembly\_fragment:00745  
fragment\_chain:1"  
17843..25453  
/note="assembly\_fragment:02138  
fragment\_chain:1"  
25554..29396  
/note="assembly\_fragment:02196  
fragment\_chain:1"  
29497..36990  
/note="assembly\_fragment:02037  
fragment\_chain:1"  
37091..57180  
/note="assembly\_fragment:01357  
fragment\_chain:1"  
57281..70963  
/note="assembly\_fragment:01309  
fragment\_chain:1"  
71064..78142  
/note="assembly\_fragment:01041  
fragment\_chain:1"  
78243..84646  
/note="assembly\_fragment:01830  
fragment\_chain:1"  
84747..104953  
/note="assembly\_fragment:02018

fragment\_chain:2"  
105054..129704  
/note="assembly\_fragment:00176  
fragment\_chain:2"  
129805..151925  
/note="assembly\_fragment:00349"  
152026..160656  
/note="assembly\_fragment:02246"  
BASE COUNT 36500 a 43576 c 42966 g 36505 t 1109 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 160656  
Score: 4378.00 Matches: 1126  
Percent Similarity: 36.79% Conservative: 296  
Best Local Similarity: 29.13% Mismatches: 425  
Query Match: 43.82% Indels: 2025  
DB: 2 Gaps: 47  
US-09-964-956-13 (1-1896) x AL592072 (1-160656)  
QY 1 MetLysAlaMetProTrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMet 20  
Db 55366 CTGCCGGCCATGCC-----TCTGTCTGCTCTCTCTCTG---CTGCTCTTCCCTGCCGTG 55316  
QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40  
Db 55315 GGGGGGGCC-----CTGGGCAACAGGCCCTTC-----CGTGCCTTC 55280  
QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60  
Db 55279 GTGGTG-----ACAGACACACAGCCTTACCCACCTGTGTGACCGGTGACT 55232  
QY 61 GlyHisIleTyrluGluAlaValAsnArgIleTyrluLysLeuSerSerAspLeuLysVal 80  
Db 55231 GGGGAGGTGTTCTGGGCGCAGTGAACCGAGTCTTTAAGCTGCCCCACCTGACTGAG 55172  
QY 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrluProArgIle 100  
Db 55171 CTGGGGCCCATGTCTACGGGGCCCGTCGAGGACACGCTCGCTACCGCCCCCAGC 55112  
QY 101 ValGlnThrCysAsnGluProLeuThrThrAsnAsnValAsnLysMetLeuLeu 120  
Db 55111 ATCGCGCTGTGTGCCACACCGCTGCGCCCCCGTGCACCAACACAGCTGCTGCATATA 55052  
QY 121 AspTyrluLysGluAsnArgLeuIleAlaCysGlySerLeuTyrluGlnGlyIleCysLysLeu 140  
Db 55051 GACTATCGCGCGCGCCCTGTGGTGGCTGCGGCGAGCATCTGGCAGGCATCTGCCAGTTC 54992  
QY 141 LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrluHisLysLysGluHisTyrlu 160  
Db 54991 CTGGCTCTGGAGCACCTTCAAGCTGGGTGAGCCGCCACACCGAAGGAGGACTACCTG 54932  
QY 161 SerGlyValAsnGluSerGlySerValPheGlyValIleValSerTyrluSerAsnLeuAsp 180  
Db 54931 TCGGGGCGCCAGGAGCGCCGACTCCATCGGTGTGTGTCATTTGTGAGCAGGCGCCAGG 54872  
QY 181 AspLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrluPheProThrIleSer 200  
Db 54871 ACCAAGCTGTTTGTGGGCGACTGCTGTGCGAGCGGCAAGTCGGAGTACTTCCACCTT 54812  
QY 201 SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAla----- 215  
Db 54811 TCCCGCAAGCTCATCAGTGTGTAAGACAGACGCGGCGGACATGTTTTCAGTCTCGTGG 54752  
QY 215 ----- 215  
Db 54751 TCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 54692  
QY 215 ----- 215  
Db 54691 TCAGAGGGGACTTTTCGGCTCTCTCAGTGTCTGGGATGCAGACAGCTTTCGGGAGGTGG 54632

Qy	215	-----	215	Db	53551	ACCCTAGGCGAGCAGTCAGTCTGCTGGTGGAGGGACCGGGTCTCTAGATCCCGCTCTCTT	53492
Db	54631	GACAGCCTGAACCCAGGTTGGGGGGTCCCTGTGTGAGGAGGCTGGTCACTCGCC	54572	Qy	399	-----	399
Qy	215	-----	215	Db	53491	CTGGGACTCAGGAATGCGCGGCTCCATCTTGGCCCTGGGAGCTTTGACTCTCTACAGGTT	53432
Db	54571	TCTGAGCCTTTCTGGCTGGGCTCTGTGTATCATCCAGGGGAGGGGATGCAGAGG	54512	Qy	400	-----	410
Qy	215	-----	215	Db	53431	CCTCCTCTCTTTCCACGAGCCCATGCAGATCAACGGCAACTCTGTGGCTGGTGTGAAC	53372
Db	54511	AAGCTGGGGTGGCCGTGATGTGGACACAGGGCTCACCGGATGCTCTCTCTCCCT	54452	Qy	411	AlaProLeuGlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAsp	430
Qy	216	-----TyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr	233	Db	53371	CAGCCTCTGGGAGGCTGTCATGTGATCGAGGGCTGCCCTGTGTGGCCGACAGCAGCAGAC	53312
Db	54451	GCTCCCGAGGTACACAGAGTGTGTGTCTCCACAGATCAAGATCCCTCTCAGACAG	54392	Qy	431	ArgMetThrSerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThr	450
Qy	234	PheThrIleLeuProAspPheAspIleTyrValTyrGlyPheSerSerClyAsnPhe	253	Db	53311	GGCATGCCAGCGTGGCCCTACACCTACCGCCAGCACCTCTGTGGTCTTCATTTGGCAGC	53252
Db	54391	CTGCTCTGTACCTGCTTTGACATCTACTACATCTACGGCTTCGTCAGCGCCCTCTTC	54332	Qy	451	LysSerGlyLysLeuLysLys	457
Qy	254	ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu	273	Db	53251	CGCAGCGCAGCTTGAAGAAGGTGGCCCCCAGAGCCCTGGGCTGTGGGTGGGGACAG	53192
Db	54331	GTGTACTCTTCAGCGCTGCACCTGGACACCCAGCAGAGCGTGTGGACACAGCGGGCAG	54272	Qy	457	-----	457
Qy	274	GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr	293	Db	53191	TCTAGATATGGACAAGGCTGCTGTGGGAACACACAGGGAGAGCTCTGAGGACAGGACA	53132
Db	54271	AAATTTTTCACGTCACAGATCGTGGCATGTGCGGGGAGACTCAGAGTTCTACTCATAC	54212	Qy	457	-----	457
Qy	294	ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla	313	Db	53131	GGAGAGACACGCTGGTGCAGGAGGGAGATTTCCCTGGCTGCCAGGCCCTGGC	53072
Db	54211	GTGAAATTCCTCATCGCTGCTCTGGCGCGGTGGAGTACCGCTTGTGTGCAGAGCGCC	54152	Qy	457	-----	457
Qy	314	TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspAsp	333	Db	53071	CCTGCTCCTCGGCTCGCCCTTGGCGCCACCCCTCACCATTTGCCAGTTCCCACTGTGA	53012
Db	54151	CACCTGGCCAGCCTGGCTGCTGCTGCCAGGCCCTGGGGTCCCGCTGATGAGGAC	54092	Qy	457	-----	457
Qy	334	LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer	353	Db	53011	CTGAATTCAGCTTCGTATGGCTCTGGGACAGAGAACACACTATTGTGGAGTCTCCC	52952
Db	54091	GTCTCTTTCACCATCTTCTCAGGGCCAGAGAACCGGGCCAGCCACCCCGGAGACC	54032	Qy	457	-----	457
Qy	354	AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer	373	Db	52951	TGCCACGAGACGAGCCTGGGGCTGAGCAGCATTTGCTGGTCTCTGTTGAATGAACAA	52892
Db	54031	ATCTCTGCTCTTACCTCCTCAGCACATCAATGCCACATCCGGGGCGGATCCAGTCC	53972	Qy	457	-----	457
Qy	374	CysTyrArgGlyGluGlyThrLeuAspLeuAlaThrLeuLysValLysAspIleProCys	393	Db	52891	GGGAGCAAAATGCCTTGGCCCTCAAGCCTCCCTATGCTGGGACAGACCCACTTGCCTCG	52832
Db	53971	TGCTATCTGGGAGGGCACTCTGGCTCTGCCCTGGCTGCTGAACAAGGAGTGCCTGC	53912	Qy	457	-----	457
Qy	393	-----	393	Db	52831	CCCTCTCTTGGCTCACTACTTCCCATCTTCCCTCTACCCATGGGCCCGGGGGCTGA	52772
Db	53911	ATCAACACCGTAGCCCTCATACCCACACACTGTGCTCTGCCCTGTCCAGGTCTACC	53852	Qy	457	-----	457
Qy	393	-----	393	Db	52771	AGCTGCTTCCATTTGGGATCTGGGGTAGCTTGGCATGAAGGGGAGATGAATCTCTGG	52712
Db	53851	ATGCCACGCTGGGCTCAGGCCAGTCACTCTGTCTCCAGGCTTTGCCATGGCCCTGGAG	53792	Qy	457	-----	457
Qy	393	-----	393	Db	52711	CATCAGGTACGTGTTGAATAGGTAGGATCGGGGACTCGGGGACTCGAGTCTCTCATCTGTG	52652
Db	53791	TGGCACCTTTGCCCTCCACCATGTTTCTCCCTCCAGGGTCCCTGCCCTCTCTCTGCT	53732	Qy	457	-----	457
Qy	394	-----SerSerAlaLeuLeu	398	Db	52651	CCTCGTAGATGCAGTGCCTATCCCGGAGGGCAAGTGTGGTCTGGGGCGCAGGAG	52592
Db	53731	CTCTCTGGGCTGCCAGTGCCTCACTTCTGTCTTCTGCTTCCGTTCCAGAGCTCTCTCT	53672	Qy	457	-----	457
Qy	399	Thr	399	Db	52591	GGTTTGTGGAGAAGGATGATGAATCTCTCTGGCCACCTGCTGTGTCTTCCCTGGGACCTCT	52532
Db	53671	GCTCGTCTCCGGTGGCCCTTGTCTTGAGCCACACAGAGTCTTGTGGGCCAGGCTTC	53612	Qy	457	-----	457
Qy	399	-----	399	Db	52531	TTTCTCTTGTGAGAGCTTTGTTGAACCGCTACTATCAATTCACCCCATTTTAAAGTGAC	52472
Db	53611	GCTCTGCT	53552	Qy	457	-----	457
Qy	399	-----	399				

Db 52471 AGTTGAGTGGGTTCTGTGGATCTGCAGAGTGGGACACCAACATGGCCACCTACCTCATTTCTC 52412  
QY 457 ----- 457  
Db 52411 CTAATCAATGCCATCTTCTAGCTCTTCCTCCCTCCCTGCATCCCTGACCTTGGGCAACAC 52352  
QY 457 ----- 457  
Db 52351 GAATCACTCCGCTTCTGTGCTGAGGGCATTTTCTGGCGTCTGCACCCCTCTTTT 52292  
QY 457 ----- 457  
Db 52291 TGGCCTCCGGGCTGTGGGTGTGAGTCTGAACCCCAACAGGTCCTGACGTACGCTCAGCC 52232  
QY 457 ----- 457  
Db 52231 ACAACCTCAAAGTCTTTGGCTGGAGAGACAGAGTGTCCCAGGTGCAGAGGCTGTGGTG 52172  
QY 458 ----- IleArgValAspGlyProArgGlyAsnAlaLeuGln 469  
Db 52171 CATCAGGCTGGTCTGTGGCTCAGGTGGCGGTGATGGCTCCAG--GATGCCCACTG 52115  
QY 470 TyrGluThrValGlnValValAspProGlyProValLeuArgAspMetAlaPheSerLys 489  
Db 52114 TATGAGACAGTCCCGTGTGTGATGACGACGCCCATCTCCGAGACCTGCTCTTCAGCCCG 52055  
QY 490 AspHisGluGlnLeuTyrIleMetSerGluArg----- 500  
Db 52054 GACCACCGCACATCTATCTCTGAGTGAGAACAGAGTGGGCTGTGGTGGCGGTG 51995  
QY 501 ----- GlnLeuThrArgValPr 506  
Db 51994 GCTGTGGGGCGGGGTGGGCTGGGTGCTGATGGCTGTCCCCAGGTGAGCCAGCTCC 51935  
QY 506 oValGluSerCysGlyGlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHi 526  
Db 51934 GTGGAGACCTGTGACCACTACAGAGCTGCGGAGCTGCTGGGCTCCGGGACCCGCA 51875  
QY 526 sCysGlyTrpCysValLeuHisAsnThr----- 535  
Db 51874 CTGTGTTGTGTCTGCTCGACACAG -GTAGGGCGGGGGCCCTGCTCGGGGAGTTGG 51816  
QY 535 ----- 535  
Db 51815 AGGGCCCCACTGGCCAGGGGGAGCCAGCCACCCAGCCGTCCTTGGGGCTCCGCC 51756  
QY 535 ----- 535  
Db 51755 GCCCTCTCCACTTCTCCAGCTAATGAGTGGAAACCTCCACCCGAGTGACGTCCCTCGG 51696  
QY 535 ----- 535  
Db 51695 GCCCAGGGGACGCGCAAGGCTGCTGTGGCTGGCACTGTGAGCTGGACAGGCTGG 51636  
QY 536 ----- CysThrArgLy 539  
Db 51635 GCCCTCCCGGGCAGTGGCGGGACCGGCTCTGGCCGACCCCGTGCAGGTGCTCGCCGCA 51576  
QY 539 sGluArgCysGluArgSerLysGluProArgPheAlaSerGluMetLysGlnCysVa 559  
Db 51575 AGGGGCTGTCTGGGGCCCTCTGCCACACAGGCTTTGTGAGAGAGTCTGAGCAAGTGTGT 51516  
QY 559 lArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuLeu----- 577  
Db 51515 CCAGGTGCGGGTCCGCGCCCAACAATGTGTCAGTGACGTCACTCCCTGGGGTGCAGGTGACAG 51456  
QY 578 ----- ValLeu-GluThr----- 581  
Db 51455 CTTGGGGGTGCCGGGTGGGTGTGCATGTGTGCTGGAGTCCCGCCCTGCCCTGAGCC 51396  
QY 582 ----- TyrAsnValProGluLeuSerAlaGlyVala 592  
Db 51395 TCTGTCTCCCGCAGCTGACCGTCAACCGTGCAGACCTTCAGTGGCGGGGTGA 51336

QY 592 snCysThrPheGluAspLeuSerGluMetAspGly---LeuValValGlyAsnGlnIleG 611  
Db 51335 GCTCGGCTTCGAGCGCGGCGGAGAACGAGGCGGTCTCTGCTGCCCTCCGCTGAACGTC 51276  
QY 611 lncysTyrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAsp---- 629  
Db 51275 TCTGCCCTCACCTCCCTCCAGGAGTCCGAGCTCTTACCAGGGGCGATGTCAGTGG 51216  
QY 629 ----- 629  
Db 51215 TTGGGGCTGCCAGATGGGCGACAGTGGGGCTCTCCCTACCCCGAGGAGTTCACGGC 51156  
QY 630 -----His-ValValGlnLeuGlnLeuLysSerLysGlu 641  
Db 51155 CACCCCGGAGTGTGTCAGGGGCCACCGCACTGTGCGGCTGCACGCTCTCTCCAAGGAG 51096  
QY 642 ThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerValHisAsnSer 661  
Db 51095 ACAGCGGTGAGTTTGGCGGTGCTGACTTTGTCTTCTACACTGCAGGCTCTCCAGTGG 51036  
QY 661 ----- 661  
Db 51035 TGAGTACCTGCCAGCACCCCTCCCTACCCCTGGGATAGAGGGGACCTCTCGGACAGGC 50976  
QY 661 ----- 661  
Db 50975 GACACCTTCAGGGCAGCTGTTTCTGGGGCATCAGGGGCTAACTGTCAAGCCCTGATAGC 50916  
QY 661 ----- 661  
Db 50915 CCTGACATCCCACTATCTCTGTCTCCCTGATATGGCTGCTCTTCTTCTTCTTCTTCT 50856  
QY 661 ----- 661  
Db 50855 TGTCTCTATTGGGGAGCAGTGAAGGGCAGTGCAGCCCTCCTCTGTATTCTTGAACCCAC 50796  
QY 662 -----CysLeuSerCysValGluLeuProTyrArgCysHisTrpCysLysTrp 677  
Db 50795 TTCCCGCCCGAGTGTCTGTGTTGGCAGCCCTTACCCTGCTCACTGGTGAAGTAC 50736  
QY 678 ArgHisValCysThrHisAspProLysThrCysSerPheGlnGluGlyArgValLysLeu 697  
Db 50735 CGCCACAGCTGTACAGCCGCCCGCCCGCCAGTGTCTCTTCCAGGAGGGGCTCCACAGC 50676  
QY 698 ProGluAsp----- 700  
Db 50675 CCTGAGGT-GAGGGCGGGCGCCATGTGAGGGGCTGGGCTGTGTGGTGGCGGGGGCCA 50617  
QY 700 ----- 700  
Db 50616 CCGGCTCTATCGGTTCTCGGTTTCTTTGAGCCTTCTCCAGCTTGGGCTGGAGAAGGT 50557  
QY 701 -----CysProGlnLeuLeuArgVal 707  
Db 50556 GGCCTCTGTCCAGCTGACAGTGTCTTCCCGCAGGCTGCCCTGAGATCTCGCCACT 50497  
QY 708 AspLysIleLeuValProValGluValIleLysProIleThrLeuLysAlaLysAsnLeu 727  
Db 50496 GGGCACTCTGTATCCCGCTGGGGTGTATGAGTGTGCGGTGTGCGGCTTGAAGAACCTA 50437  
QY 728 ProGlnProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGlu 747  
Db 50436 CCTCAGCGCAGTCCGGGCGAGAAACTATGATGCGTGGTGGGCTGCGGGTGCAGGGCGGCGAG 50377  
QY 748 GlnArgValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThrSer--- 766  
Db 50376 CAGCGGTGCTCGCGTGCCTTCAACAGCAGCAGTGTGCAGTGCACAGAACCCCTCGGTG 50317  
QY 766 ----- 766  
Db 50316 AGTCCCCACCGCTGCCCTCCCTTCCGGGGTGTGGGCTGTGGGCTGTGGTGTGGATCTTCTT 50257

```
QY 766 ----- 766
Db 50256 TCTATGCCAGCTCTTGGCAGGAAGTGTGAGTCTCTGCTAGGATTCCTGTACCT 50197
QY 766 ----- 766
Db 50196 GGAGGAGGGGACGTGGCAAAATCTTGGTAGGGTTCTGCACATCTTTATCTGGGGT 50137
QY 767 -----
Db 50136 CCATGGGTTCTTCTCCAGTACTCTATGAAGGTGATGAGCATGTGCACCGAGCT 50077
QY 779 lGluLeuThrValValTTPAsnGlyHisPheAsnIleAspAsnProAlaGln-AsnLysV 799
Db 50076 GGACTTCTCGTGTCTGGGATGAGACTTCCCATAGACAAGCTCCACGTTCCGAGG 50017
QY 799 alHis----- 800
Db 50016 TGAAGCATGGGCAGGGAGCTTCCCTCCTAAGGCAATTGGCACTGCTGGGGTCTCGGTCT 49957
QY 800 ----- 800
Db 49956 GGGGGCCCTAGTGTCTCCCACTTTCCACTTTTCTGCCACTGCTGCTCCGTCAGCAGTGC 49897
QY 801 -----
Db 49896 CTTCTGTGCTGCAGCCCTCTCTGTACAAAGTCTGGGCGCAGCGCCAGCTGTGGCCCT 49837
QY 814 ysLeuLysAlaAspProAspPheAlaCysGlyTTPCysGlnGlyProGlyGlnCysThrL 834
Db 49836 GCCCAAGGCTGATCCCGCTTCAACTGTGCTGTGTCATCTCAGAGCACAGGTGCCAGC 49777
QY 834 euArgGlnHisCysProAlaGlnGlnSerGlnTTPLeuGluLeuSerGlyAlaLysSerL 854
Db 49776 TCCGAGCCACTGCCCGCCCGGAGCAACCACTGGATGACCTGAGCAGCAGGCGCACCC 49717
QY 854 yscysThrAsnProArgIleThr----- 861
Db 49716 GGTGAGCACCACCCCGCATCAGCAGGTCAGCCTCCCTCAGCGCCCTGCCACATGCCAA 49657
QY 862 -----
Db 49656 CAGGCGCCCTGGGAGCTGAGCCAACTCTCAGTGCCTCCTGCTCCACAGATCCACC 49597
QY 865 roValThrProArgGlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyL 885
Db 49596 CTCTCGTGGGCGCCCAAGAGGAGCACCAGGTCACCATCTGGGTGACAACTGGGCC 49537
QY 885 euGluPheArgAspIleAlaSerHisValLysValAlaGlyValGluCysSerProLeuV 905
Db 49536 TCTTGTCCGAGAGGTGGGC-----CTGCGGGTGGCTGGCGTGGCACTCCATTC 49483
QY 905 alAspGlyTyrIleProAlaGlu----- 912
Db 49482 CGGCGGAGTACATAGTCTGAGAGGTGAGTGGGCTCTGTGGGTGCGCGGCCGTATGT 49423
QY 913 -----
Db 49422 GGCCTGGCGGCCCTGACGCTCTCTGAGCCCTAGGATCGTGTGAGATGGAGAGTCGC 49363
QY 922 --LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgProGluP 941
Db 49362 TGGTGCCCGCCCGCGCGCGGCGCGGTGGAGCTGTGTGGGTGACTGTTCAGCGGACT 49303
QY 941 heMetAlaArgSerSerGlnLeuTyrTyrPheMet----- 952
Db 49302 TCCGACCCAGTGGAGGAGGTCTACAGCTTTGTGGTGGCTGCTGGCGGCCCTACCCCT 49243
QY 952 -----
Db 49242 TCCTGTCCCTCTCTCTCCGCAAGGGGCGTGGAGCAGCGCGCGCGCTCTCCCTCCCT 49183
QY 953 -----
Db 49182 CAGGCGAGCTTCTCCCGCAGACCCCAACGTTTGGACCAAGTGTAGTCCACCGCTGGCCCGG 49123
QY 956 etSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnValV 986
Db 49122 CGTCCGGGGCACACAGGCTTACCATCTCAGCAGCTCTCTGGATGTGGCAGCAGGTCA 49063
QY 986 alValMetPheGlyLysGlnProCysLeuPheHisArg----- 998
Db 49062 CAGTCACTGTGAGGACACCGAGTGCAGTTTGTAAAGTGGGCGCGGCCCTGCCAGCTT 49003
QY 999 -----
Db 49002 TGGGTTGGCATCTGTGGGGCGCGTGGGACGGGTGGCTGAGGGCCCTGGGCGCACCG 48943
QY 1000 erProSerTyrIleValCysAsnThrThrSerSerAspGluValLeu----- 1015
Db 48942 CTCCAAGCACCTCTGCT-TGCCAATGTAGGAGATGCCAAGGCGATCGTGTGATCTCAC 48884
QY 1016 -----
Db 48883 CTCTCTCCACCTGGGCCCGCAGCCAGCCCGCCCATCACACTTGCATTGACCGGCTAA 48824
QY 1028 leHisGln---AspLeuValPheGlnTyrValGluAspProThrIleValArgIleGluP 1047
Db 48823 TCTCCAGCGCCCGGCTCATCTACACTACACTCAGGACCCACCGCTACCGCTTGGC 48764
QY 1047 roGluTrpSerIleVal----- 1052
Db 48763 CCACCTGGAGCATCAT-CAAGTAAGACCTGGGGGACTGGGGAGCCTGGCAGTGTCCAGA 48705
QY 1052 -----
Db 48704 GGGCTCAGGAGCTGGGTGCTCTGAGTCCGCTGGGGCCCTCTCTGGAGCCTAGGCCCT 48645
QY 1052 -----
Db 48644 CATGTGTGAGGGGTGGAGTGTCTGTGCTGCACAGTACCCGCGCACCCTAGGCGATC 48585
QY 1052 -----
Db 48584 CAGGTCAGGAAGGCTGGGCTGCCATGCGACGCTTGATCGCTCTCCAGGCTGGGTGG 48525
QY 1053 -----
Db 48524 GTGCACCTGGAGGAGGAGCCTGAGGCCCTGCCCTGTCCCTAGTGAAGCACTGCCATCA 48465
QY 1059 laValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGlyG 1079
Db 48464 CTGTGAGTGGGACCCACCTGCTGACGCTCCAGGAGCCCGGGTCCGTGCCAAGTACC 48405
QY 1079 lyLysGluHisIleAsnIle----- 1085
Db 48404 GCATTGAGACCAACCAATGTAGTACCAGTGCCTGCCCGCCCGCCACCCGACCCCTG 48345
QY 1086 -----
Db 48344 GGCTTCACCTGCTGCTGCCACCTTTGTCCTCCACAGATGCCAAGTATCAACGACA 48285
QY 1092 hrGluMetThrCysGlnAlaProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuT 1112
Db 48284 CTGCCATGTGTGTAGGCGCCCGGATCTTCTTGGGGGCCCGCCAGCTCGGGCGCAAG 48225
QY 1112 hrGluArgProGluGluPheGlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeuA 1132
Db 48224 GCGAGCACCTGTAGTGTGGTTCCTGCTGGACCACTGCAAAACGCGCCGCTCCCTCA 48165
QY 1132 snLysThrAsnPheThrTyrTyrProAsnProValPheGluAlaPheGlyProSerGlyI 1152
Db 48164 ACCGCTCCCTTACTACTACTACCTGATCCAGCTTGGGCGCTGGGGCCCTCTGGCG 48105
QY 1152 leLeuGluLeuLysProGlyThrProIleIleLeuLys----- 1164
```

```
Db 48104 TGGTGCACGTCAAACCGGGCTCCACACGTGTGTGTGAAGGTGCGGGGGGGTGGGGGGCGG 48045
QY 1165 -----GlyLysAsn 1167
Db 48044 GAGGGCGGGAAGTAGAGAGTCTGGGCTGAAGTTGTCTCCACCCCGAGGGCAAGAAC 47985
QY 1168 LeuIleProValAlaGlyGlyAsnValLysLeuAsnTyrThrValLeuValGlyGlu 1187
Db 47984 CTGATTCCCGCGGCGACCGGCACTCC--CGCCTCAACTACACTGTGCTGATAGGAGGC 47928
QY 1188 LysProCysThrValThrValSerAspValGlnLeuLeuCysGluSerProAsnLeuIle 1207
Db 47927 CAGCCGTGTTCGCTCACTGCTCGGACACAACTCTCTGCGACTCACCCAGCCAGACT 47868
QY 1208 GlyArgHisLysValMet-----1213
Db 47867 GGCCGCGAGCCTGTCATGTAGTGGGATGGGAGACCCCTGGGCGAGCCAGGGTGG 47808
QY 1214 -----AlaArgValGlyGlyMetGluTyr 1221
Db 47807 CGTGGTGGTCAGCTCACCTCAGCCCTGTCCCCACAGGTGTGTTGGGTGGCTGGAGTTC 47748
QY 1222 SerProGlyMetValTyrIleAlaProAspSerProLeuSerLeuProAlaIleValSer 1241
Db 47747 TGGCTGGGACCCCTGCACATCTCGGACAGCGGCGCTGACCTACCGGCATGATGGG 47688
QY 1242 IleAlaValAlaGlyGlyLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrLys 1261
Db 47687 CTGGCGGCGGGGTGGCTCTGCTGCTGCCATCACACCCGTGCTGGTGGGTACAAG 47628
QY 1262 ArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeu 1281
Db 47627 CGCAAGACTCAGGACGCGGACCGTCACTCCCTCAAGCGTGTGAGCTGCAGATGGACAACCTG 47568
QY 1282 GluSerArgValAlaLeuGluCysLysLysGluAla-----1292
Db 47567 GAGTCCCGTGTGCCCTGGAGTGCAGGAAGG--TGCCGTAGGCGGGCGGGGATGTGGTGT 47509
QY 1293 -----PheAla 1294
Db 47508 GGAAGCTGGGACCTCCCTCTGCCACCTCATTCCTCCACCCCGCCAGCTTTTGCA 47449
QY 1295 GluLeuGlnThrAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPhe 1314
Db 47448 GAGCTCAGACGGACATCAATGAGCTGCACTTAACACATGACGAGGTGCAGATCCCTTC 47389
QY 1315 LeuAspTyrArgThrTyrThrMetArgValLeuPheProGlyIleGluAspProVal 1334
Db 47388 CTGGACTACCGGACTTACGCGGTGCGGTGCTCTTCCCGGCATCGAGGCCACCCCGTG 47329
QY 1335 LeuArgAspLeuGluVal-----ProGlyTyr-----1343
Db 47328 CTCAAGGAGTGGATGTAGCCCTCTGCTGCGCTGCGCCCTCCACCATTCCTCCAGGCGCG 47269
QY 1344 -----ArgGlnGluArgValGluLysGlyLeu 1352
Db 47268 CCCCCACCTCTCAGACCTCTCTGCTCCCCACAGACGCCACCCACAGCTGGAGAGGCCCTG 47209
QY 1353 LysLeuPheAlaGlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeu 1372
Db 47208 CGCCTCTCGGCGAGTGTGTCACAGCGCGCGCTTGTGTTTACCTTCATCCACACCGCTG 47149
QY 1373 GluSerGlnArgSerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThr 1392
Db 47148 GAGGCCACAGACGACTCTCATGCGCGACCGCGCACCGCTGCGCTCGCTCACCATGGTG 47089
QY 1393 ValLeuGlnSerLysLeuGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeu 1412
Db 47088 GCCCTCAGAGCGCGGTGCACTATGCCACGGGCTGCTCAAGCAACTGCTGGCGGCACTC 47029
QY 1413 IleAspLysAsnLeuGluSerLysAsnHisProLysLeuLeuLeuArg-----1428
Db 47028 ATCGAAGAACCTCGAGAGCAAGAACACCCCAAGCTGCTGCTACCGACGTTACCTGCCT 46969
```

---

```
QY 1428 -----1428
Db 46968 CGCTCTATCAGGCCACACCTTTGTCTGCTGGCTGCCCTGTCCCAAGCCCCACCC 46909
QY 1429 -----ArgThrGluSerValAlaGluLysMetLe 1438
Db 46908 CTGCCAGCCCGAGCCCTTCTCTTGGCTTAGGACAGAGTCAGTGGCTGAGAAGATGCT 46849
QY 1438 uThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLys-----1451
Db 46848 TACCAACTGGTTCAGCTTCTGCTGCATAAAGTTTCTGAAGGTGGCCAGGTGGTGGGCG 46789
QY 1451 -----1451
Db 46788 CAGGGAGGTGGTGGCAGAGGACCGCCAGTGGTCAGGCAGGCAGCTGTGACCTG 46729
QY 1452 -----GluCysAlaGlyGluProLeuPheSerLeuPheCysAlaIleLysG 1467
Db 46728 TCCCTGGCTGAGGAGTGTGCTGGGAGCCCTCTTCTCTGCTTTACTGTGCCATCAAGCA 46669
QY 1467 nGlnMetGluLysGlyProIleAspAlaIleThrGlyGluAlaArgTyrSerLeuSerG 1487
Db 46668 GCAGATGAGAGAGGCCCCCANTGATGCCATCACGGGCGAGGCACGATACTCCTGAGCGCA 46609
QY 1487 uAspLysLeuIleArgGlnGlnIleAspTyrLysThrLeuVal-----1501
Db 46508 GGACAAGCTCATCCGTACAGCATCGACTACAAGACACTGGTGAAGCGCAGGCAGCGCG 46549
QY 1501 -----1501
Db 46548 GCCAGAGTAGGGGGCGCAGAGAGCGCTCGCCCCAGACTGACACTGGAGTCCGCTTTC 46489
QY 1502 -----LeuSerCysValSerProAspAsnAlaAsnSerProGluValProVal 1518
Db 46488 CCTCAGACCTTCAC--TGGTGTGTCCGGAGAACGAGGCGACGCCAGCTCCAGTGAA 46430
QY 1518 sIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePhe 1538
Db 46429 GGTTCACACTGTGACAGCATCACCCAGGCCCAAGATAAGCTGCTGGACACTGTGTACAA 46370
QY 1538 sAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeu-----1553
Db 46369 GGGCATTCGCTACTCCCGCGCTCCCAAGCTCAGGACACTGGACCTGGGTGAGGTCCCCAC 46310
QY 1553 -----1553
Db 46309 CCTCTCTCTGGCTCCCACTCATACCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 46250
QY 1553 -----1553
Db 46249 GCTCCCACTGAACCTGTTTCTTAAGAGAGTGGCGCAGGCACCTTGGGGGTGCCAGC 46190
QY 1554 -----GluTrpArgGlnGlySerGlyAlaArgMetI 1564
Db 46189 ATAATCTTAAGGGCTGTGTGGCCCCACAGAGTGGCGCCAGCGCCGATGACTCGCATCA 46130
QY 1564 LeuGlnAspGluAspIleThrLysIleGluAsnAspTrpLysArgLeuAsnThrL 1584
Db 46129 TCCTCCAGATGAGGATGTCACCAACAGATCGATGTGACTGGAAGAGGCTCAACTCAC 46070
QY 1584 euAlaHisTyr-----1587
Db 46069 TGGCCCACTACCAAGTGAGGGTGGGGCCCATTCCTCCCAAGGCACCTGGGAGCCAG 46010
QY 1587 -----1587
Db 46009 GACCTGCTTGAAGTGCAGCAGGACACGGGAGAGTGGCCCTGGCTCCCTCGCCCTTC 45950
QY 1587 -----1587
Db 45949 TCCCTGGCTCGTGGGCTCCCTCTGCTGGGTGGTGGGTGGGTGGGGGGCGCTTGGCTTCTCA 45890
```

QY	1588	-----GlnValProAspGlySerValVal	1595	Db	44812	GCTGGCCTGGGTTTGACACTGACGGTGCCATCAAGCAAGGGCGCTGTTCGTCACGATT	44753
Db	45889	GCTTCCTGCACTGCCCCCTCTGTCTCTGGGGCTCCAGAGTGACAGCGTTCTCTGGTG	45830	QY	1741	-----	1741
QY	1596	AlaLeuValSerLysGlnValThrAlaTyrAsnAlaValAsnAsnSerThrValSerArg	1615	Db	44752	CTGGAGCTGGACTGCTGGTGGTGGGATGTGTGGTGGAGCCCTTTGCCCTCTCTGGGGT	44693
Db	45829	GCAATGGTGGCCAAACAAAGTGTCTGCCTATPAACATGGCCCACTCTCTCACCTTCACCGCG	45770	QY	1741	-----	1741
QY	1616	ThrSerAlaSerLysTyr	1621	Db	44692	CTTGCAACAGGTAGAGATCTGGTGGTCCAGGTGCCAGGTGGTGGAGCGGTGACAGAGAA	44633
Db	45769	-----TCCCTCAGCCGCTACGGTAGGTGTCTCAGTGTGGTGCCATGTGCCCTTCGAGGGA	45713	QY	1741	-----	1741
QY	1621	-----	1621	Db	44632	GCTGGGAGGGGTTGAGCCCTATTGAGAACAGAGTTTCAGCAGAGACTTTAGAGGAAAGG	44573
Db	45712	ACCCCACTTCCAAAGTGCCATCGATTCGTAGAGTGTAGAGGGGTGCGGCCACCGAGG	45653	QY	1741	-----	1741
QY	1621	-----	1621	Db	44572	CCGTTCACCCCTAAATGTGTGTCTCACCACCTCTGGACAAGATGTGAGCGGCGCGGACA	44513
Db	45652	GCAGATGGTCCACATGCTGCTGAGCTCCGGGAGAGTGGGGCAGGGGCCAGGTGGTG	45593	QY	1741	-----	1741
QY	1622	-----GluAsnMetIleArgTyrThrGlySerPro	1631	Db	44512	GGAAAGAGATGTGCACATGGGAGGGGCCCTTGGGGCTAACTAGGAGCGCTCAGCGCAC	44453
Db	45592	GCCTAAGGGTCACATGCTCTCTCCAGAGCTTGTCTCCGACCGCCAGCGCCCT	45533	QY	1742	-----CysLeuProLeuArgPheTrpValAsnMetIleLysAsnProG	1756
QY	1632	AspSerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyValLysMet	1651	Db	44452	GTCCCTCTGTGTGCCACAGCTCGCGTGGTGGTGAATGTGATCAAGAACCCGC	44393
Db	45532	CATAGCTTCCCTCAGGGGCACCCATGATTAGCCCTGACCGACAGGACACCAATTTG	45473	QY	1756	lnPheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValValAlaIag	1776
QY	1652	TrpHisLeuValLysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGlySerLys	1671	Db	44392	AGTTCGTGTTCGACATCCACAAACAGCATCAGGATGCTGCTCTCGTGTGTAGGCC	44333
Db	45472	TGGACCTTGGTGAACACAGACCATCCGACCATCGAGGGGCGCGTGGCAGCAAG	45413	QY	1776	lnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerA	1796
QY	1672	MetValSerGluIleTyrLeuThrArgLeuLeuAlaThrLys	1685	Db	44332	AGACCTTCATGTGACTCTCTGTCTACATCCGAGCACCGCTGGGAAAGACTCGCCCTCCA	44273
Db	45412	ATGGTCTCCGAGATCTACTGNACAGGCTGTGTGGCCACCAAGGTATGGCCCTGCCTCTCG	45353	QY	1796	snLysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGlu	1812
QY	1685	-----	1685	Db	44272	ACAACTGTCTACGCCAAGACATCCCAACTACAGAGCTGGGTGGA	44214
Db	45352	CCACCTTGGCCTCCGCCAGAGGTTGTCCCGCCTTACAGGGGAGGGCCATGGATCAIT	45293	QY	1812	-----	1812
QY	1685	-----	1685	Db	44213	CCGCCCTGTGTGGTGGCAGAGGCGACCTGCCCTCCCTGGCTGTGTGCGAGG	44154
Db	45292	TGCTTCCAGTGGGCGCTTCTTCGGGTCTTTGTGTGGGAACCTCTGAGCCTTAGATAAAG	45233	QY	1812	-----	1812
QY	1685	-----	1685	Db	44153	CAGGCCAGCTACAGGCGAGGAAGCCCGGGGCTGCTGGGCACTGAGGCGAGGGTTCTAG	44094
Db	45232	CCATGTCTGTGAGAGACCACTGGCCCTGTAGGAAAGCCCTCTCTTTGCCAAGCCTTT	45173	QY	1813	-----ArgTyrTyrSerAspIleG	1819
QY	1685	-----	1685	Db	44093	TTTTTGTGATGTCGCTGAGTGACATCTACTCAGGGTCCCAACAGGATATTATCGACATTG	44034
Db	45172	CTGCCTCCATCTGTCTGTCTGGGACATCCCGACCTGGCTGCCCTCATAGCTGTGACC	45113	QY	1819	lyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluGlnSerArgM	1839
QY	1686	-----GlyThrLeuGlnLysPheValAspAspLeuPheGluThrIlePh	1700	Db	44033	CAAAAGATGGATCCATCAGCAGCAGGACATGGATGGCTACCTGGTGGAGAGTCCCGCC	43974
Db	45112	TCTTGTCCCCACACAGGGGCACACTGCAGAGTTCGTGGATGACCTCTTTGAGACAGTGT	45053	QY	1839	ethisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGlyL	1859
QY	1700	eSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLe	1720	Db	43973	TCCAGCCGAGCAGCTTACGCTCTGAGTCTGCTCAGCAGCTGTATTTCTATGTCACCA	43914
Db	45052	CAGCACAGCCACCGGGGCTCGGCCCTGGCCCTGGCCATCAAGTACATGTTCGACTTCCT	44993	QY	1859	ysTyrSerGluGluIle	1867
QY	1720	uAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrpLysSe	1740	Db	43913	AGTACCGCAGGAGGTGTGTGTATCCACAGACTCCAGTTACTTGTGCTCT	43861
Db	44992	GGATGAGCGGGGACCGCCAGATCAGCGACCCCGGATGTGCGCCACACCTTGGAGAG	44933	RESULT 15			
QY	1740	rAsn	1741	AC025784/c			
Db	44932	CRACTGGTATACCCCGTGTGGGCTGCCAGCAGCCTGTCTGGAGACTGGTGGCGGAG	44873	LOCUS			
QY	1741	-----	1741	DEFINITION			
Db	44872	ACGCTGGTGGCTGTGCTGAGACCCAGGGGCTGGAGTAGCATCCAGGAGGGAATGAG	44813	Mus musculus clone RP23-142E5, *** SEQUENCING IN PROGRESS ***			
QY	1741	-----	1741	unordered pieces.			
				AC025784			
				AC025784.8 GI:13940738			
				HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.			
				Mus musculus.			
				ORGANISM			
				Mus musculus			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 193961)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Mus musculus, clone RP23-142E5  
Unpublished  
2 (bases 1 to 193961)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavsky, L., Boukhalter, B., Brown, A., Buckett, G.,  
Campopiano, A., Castle, P., DeArelano, K., Dewar, K., Diaz, J.S.,  
Collins, S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardy, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (14-NAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 4, 2001 this sequence version replaced gi:13273453.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5300

Center clone name: 142\_E\_5

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 50356: contig of 50356 bp in length  
\* 50357 50456: gap of 100 bp  
\* 50457 92748: contig of 42292 bp in length  
\* 92749 92848: gap of 100 bp  
\* 92849 94975: contig of 2127 bp in length  
\* 94976 95075: gap of 100 bp  
\* 95076 144036: contig of 48961 bp in length  
\* 144037 144136: gap of 100 bp  
\* 144137 170034: contig of 25898 bp in length  
\* 170035 170134: gap of 100 bp  
\* 170135 193961: contig of 23827 bp in length.

FEATURES  
source

1. 193961  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-142E5"  
/clone\_lib="RPCI-23 Female Mouse BAC"  
BASE COUNT 52508 a 45525 c 43881 g 51443 t 604 others  
ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	193961
Score:	4372.50	Matches:	1114
Percent Similarity:	37.72%	Conservative:	291
Best Local Similarity:	29.91%	Mismatches:	433
Query Match:	43.77%	Indels:	1899
DB:	2	Gaps:	44
US-09-964-956-13 (1-1896) x AC025784 (1-193961)			
QY 8 TrpThrCysLeuLeuSerHisLeuLeu-----MetValGly-----Met 20			
Db 117824 TGGTCATGCCCACTGCTGCG-CTTCTCCCATTTGCTATTCTTCCACATAGAGGGTGCGCTG 117766			
QY 21 GlySerSerThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPhe 40			
Db 117765 GGTAGCAGC-----AGGCCATTT 117748			
QY 41 ValThrPheArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThr 60			
Db 117747 CGTACCTTTTGGTGAACAGATACCACT---CTGACTCACCTGGCTGTGCACCGAGTGACT 117691			
QY 61 GlyHisIleTyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysVal 80			
Db 117690 GGGGAGGTGTTGTAGTGCCGCTGAATCGAGTTTTCAGCTGGCTCCAAACCTAACTGAG 117631			
QY 81 LeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIle 100			
Db 117630 CTACGGGCCCATGTACAGGGCCCATTTGAGGACAAATGCTCGCTGTACCCACCTCTAGC 117571			
QY 101 ValGlnThrCysAsnGluProLeuThrThrAsnAsnValAsnLysMetLeuLeu 120			
Db 117570 ATGCGTGTGTGCTCCACCGCTGTCCTGTGGACAATGTGAACAAGCTGCTCTCATATA 117511			
QY 121 AspTyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyLysLysLeu 140			
Db 117510 GACTATGACGCCGTCGTTGGTAGCTGTGGCAGCATCTGGCAGGCGCATCTGCCAGTTC 117451			
QY 141 LeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu 160			
Db 117450 CTGCGCTGGATGACCTCTTCAAGTTGGTGGTGGCCGCCACCATCGCAAGGAGCATACTATCTG 117391			
QY 161 SerGlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAsp 180			
Db 117390 TCAGGGGCCAGGAGCGCTGATTTCCATGGCTGGCTCATTTGTTGAACAGGTCCAGGGCGCT 117331			
QY 181 AspLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSer 200			
Db 117330 ACCAAGTTGTTGTGGGCACCGCTGTTGATGGCAAGTCTGTAGTACTTTCCCACTTGAGT 117271			
QY 201 SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPhe----- 214			
Db 117270 TCCCGTAAAGCTCATCGACGATGAGGACACATGGAGATATGTTTTCAGTCTGGTGGTGAGCTT 117211			
QY 214 ----- 214			
Db 117210 GTTGCTTTTGTCTTTTCCCTGTTTCTTATGAGCCAGGCTCTTATTCACCTCTTCTC 117151			
QY 214 ----- 214			
Db 117150 TCAGTCTGGCCCTGCTTCTTAGGAAAACATTTTGGTTTGAGAGAGGAATTTATGACCCCTTC 117091			
QY 214 ----- 214			
Db 117090 ACTGTACATAGATGTAGACCAGCTGTTGAGGACAGGAAGATGGGCGAGAGTTTACCTGTG 117031			
QY 214 ----- 214			
Db 117030 GAAAGTCTCTATATAGAGAGGCTAAGGTCACCTTGCACTCTGTGGCATGGAATCCAGTAA 116971			
QY 214 ----- 214			
Db 116970 TTGTCAGGTGGGCGAGATGGGAGATCACTGGGGAAGCCCTTTAGAAAACCGCATTTGTTGG 116911			



Db 114755 CCTGAGTGACGTCCTTGGATCTCAGGGGATGAGGCCAAGGTACACTGTTGGCTGGCAG 114696  
QY 534 ----- 534  
Db 114695 TATGAGTGGGTATGCTGTCTTCTGTCCACCTCAGCCCTTACTGCACCCCATCTGGG 114636  
QY 535 ----- ThrCysThrArgLysG 540  
Db 114635 CCCTCTGGGACAGTGTCTGGACTGCTCTGGCCTGTCTCCATGCAGGTGCTGCGTGAAG 114576  
QY 540 luArgCysGluArgSerLysGluProArgPheAlaSerGluMetLysGlnCysVala 560  
Db 114575 GGGCTCTCCAGGTGCTCTGCCCCAGCTGGCTTTTCAGAAAGACGTGACCAATGTATAC 114516  
QY 560 rgLeuThrValHisProAsnAsnIleSerValSerGlnTyraSerValLeu----- 576  
Db 114515 AGGTGGGTGCTCCCATATATGTGTAGTGACATCCCTCTGGGTGCA-GGTAAGTGAC 114457  
QY 576 ----- 576  
Db 114456 ATAGAGATPACTTGGCCCTGTATATGCTTCATGAGGGAGGTGCACCCCTGCTCTGAACCC 114397  
QY 577 ----- LeuValLeuGluThrTyraSerValProGluLeuSerAlaGlyVala 592  
Db 114396 TCTGCTTCCAGCTGACCGGTAGCCATCGCAACCTGCCAGACCTCACTGTGGGTGA 114337  
QY 592 snCysThrPheGluAspLeuSerGluMetAspGlyLeuValVal---GlyAsnGlnIleG 611  
Db 114336 GCTGTCTCTTGGAGGTGACTGAAGTGAGGCTATTCTGCGCCCTCTGGAGAGTTAC 114277  
QY 611 InCysTyraSerProAlaLysGluValProArgIleIleThrGluAsnGlyAsp--- 629  
Db 114276 GATGCTCTTACCATCCCTTCCAGGACTCCAGACACTTACCAGAGGCGATGGTCACTGGG 114217  
QY 629 ----- 629  
Db 114216 ATGGGGCAATCGGGGTGTGTGGGCGCAGGCGAGAAATCCAGACTAGTTGGCAATAGCCA 114157  
QY 630 ----- HisHisValValGlnLeuGlnLeuLysSerLysGluThrG 643  
Db 114156 TCTGGCTTCTGCAGGGGCACTATACACTGTCGGGCTTCCAGTCTCCATGAGACTG 114097  
QY 643 lyMetThrPheAlaSerThrSerPheValPheTyraSerCysSer----- 657  
Db 114096 GTGTGAGGTTTGTGGGGTTGACTTTGTCTTCTTACAACTGCAGTGCCTCCAGTCGTGAG 114037  
QY 657 ----- 657  
Db 114036 TACCTGGTCCAGAACTGTCCCTAGCCCTGGGATAGTAGTAAGTGTCTGAAATATAT 113977  
QY 657 ----- 657  
Db 113976 AAAAATTGCACGTAACTTTCTCTAGGGTTACAGGGGTGAGTTGTTGGGACTTCCCTGCTGT 113917  
QY 658 -----V 658  
Db 113916 CGCCACTTCTGTCTAGTCTGATGTGCCCTATGTGGCTTGGGAGGTTCTGCCCTACCCCTA 113857  
QY 658 alHis----- 659  
Db 113856 TTCATTTGTCCCAATATGTCTTGTGGGTGGGTGATACAGGTGCGTATGACCTCTGTTA 113797  
QY 660 ----- AsnSerLysLeuSerCysValGluSerProTyraArgCysH 673  
Db 113796 TTTCTAAAGCTACTCCCCCTCCAGGTGTATGTCCTGTGTGGCAGCCCTTACCCCTGCC 113737  
QY 673 isTrpCysLysTyraArgHisValCysThrHisAspProLysThrCysSerPheGlnGluG 693  
Db 113736 ACTGTGTAACTACCTCATGTGTGTACCAGCCACCCACAGAGTGTCTTCCAGGAGG 113677  
QY 693 lyArgValLysLeuProGluAsp----- 700  
Db 113676 GCAGGTCCACAGCCCTGAGGT-GAGGTGGTGCAGTGCACCCGTGGAAGAGGCTGGGCTCC 113618

QY 700 ----- 700  
Db 113617 TTGGGGGCTGCCCTCTCACATTTTTTTGTTGACCTCCCATTTCCATGGCCCTGGAGAAAG 113558  
QY 701 -----CysProGlnLeuLeuA 706  
Db 113557 CGGTAGACCAAGTGTCTTAACAAGTGTCTCTCCAGGGCTGCCCTGAGATCTGCTGC 113498  
QY 706 rgValAspLysIleLeuValProValGluValIleLysProIleThrLeuLysAlaLysA 726  
Db 113497 CTCAAGGGAGCTCTTGATTCCTGGGTGTCATCAGCCTCTAACTCTCGCGGGTAAAG 113438  
QY 726 snLeuProGlnProGlnSerGlyGlnArgGlyTyraGlyCysIleLeuAsnIleGlnGly 746  
Db 113437 ACCTGCCACACCTCAGTCCGACAGAGAATTATGATGCGTGGTCCGTGTACAGGAC 113378  
QY 746 erGluGlnArgValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThrS 766  
Db 113377 GGCAACATCGGGTACCTGCAGTGCCTTTTAAACAGCAGCAGTGTGCAGTCCAGATGCTCT 113318  
QY 766 er----- 766  
Db 113317 CGGTAACACCAGAGTGCAGGTGGGCGCAGCAGCAATAATTTGTTGTTGTTGTTT 113258  
QY 766 ----- 766  
Db 113257 TGTATTATTTTGTGTTTTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 113198  
QY 766 ----- 766  
Db 113197 TTTGTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 113138  
QY 766 ----- 766  
Db 113137 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 113078  
QY 767 -----TyrSerTyraGlyMetGluIleAsn 775  
Db 113077 TTATTGTGCTCATGTGGTTCCTTTCTTAGTACTTCTATGAGGTGATGAGTTGTTGTTGTT 113018  
QY 776 AsnLeuProValGluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsnPro--- 794  
Db 113017 GACACCAACTGGAGCTTCTCTGTGTTGGATGAGATTTTCCCATTTGATAAGCCTCCT 112958  
QY 794 ----- 794  
Db 112957 AGCTTCGAGGTGAGGGGTGACATGGGCGCAAGGAGCATGGCTCTCTGGGGCAGGATTGGC 112898  
QY 795 -----AlaGlnAsnLysVal 799  
Db 112897 ACTCTGGGGCAAGTCTGTAGTCTTAGAGCCCCCTGCTTTTCTGCCCAT-TGCCCTGCTC 112839  
QY 800 His-----LeuTyraLysCysGlyAlaMetArgGluSer 810  
Db 112838 CATCAGTAGTCTCTGCTGCCCTGCAGCCCTCTTTTACAAGTGTGGCTCAGCGGCTAGC 112779  
QY 811 CysGlyLeuCysLeuLysAlaAspPheAlaCysGlyTrpCysGlnGlyProGly 830  
Db 112778 TGTGGCTCTCCCTCAAGGCTGATCCCGATTTCAACTGTGGCTGTGTGATCTCAGAGCAC 112719  
QY 831 GlnCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuLeuLeuSerGly 850  
Db 112718 AGGTGCCAGTGGGCTCACTGCCAGCTCCCAAGTAAGTGTGATGATGATGATGATGATGATGAT 112659  
QY 851 AlalysSerLysCysThrAsnProArgIleThrGluIle----- 864  
Db 112658 AAGGTGCCCATGAGCCATCCCGAATCACCAGGTAGTTCACCTTACCAATCTAC 112599  
QY 864 ----- 864  
Db 112598 CTTGCTTGTAGGCAATGTGTGAGTGTGGCTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 112539

```
QY 865 -----ProValThrGlyProArgGluGlyGlyThrLysValThrIleArgGlyGluAs 882
Db 112538 GATTTCATCTACAGACCCCAAGAGGGTGGCACCGGTCCACCATTTGGGTGAGA 112479
QY 882 nLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyValGluCysSe 902
Db 112478 CCTGGGCGCTCATTCCCGTGGAGGTGGC-----CTTCGAGTAGCTGGTACCTTGCAA 112425
QY 902 rProLeuValAspGlyThrIleProAlaGlu----- 912
Db 112424 CTCCATTCCACCGCAATATGTCAGTGTGAAGGTGAGTGGCTATGTGAGCACCCCTGT 112365
QY 913 -----GlnIleValCysGluMetG1 919
Db 112364 CCTTGTTTACCACAGCTGGTCTGACCTTCTGTGACTCTTGTAGGATCGTATGTGAGATGGA 112305
QY 919 yGluAla---LysProSerGlnHisAlaGlyPheValGlnIleCysValAlaValCysAr 938
Db 112304 GGAATCGCTGGTCCAGCCACCACCTGGCCCTGTGAGTCTGTGTAGGGGATTGTTTC 112245
QY 938 gProGluPheMetAlaArgSerSerGlnLeuTyrTyrPheMet----- 952
Db 112244 TGCTGACTTCGGCACACAGTCCCGCAACTCTACAGCTTTGTGGTGGTGTCTAACTAG 112185
QY 952 ----- 952
Db 112184 CCCCTTCCCCTTCCCTATCTGGGAAATGTGGAGTGTCCCTTGGATCAGGCTGGCTTCTA 112125
QY 953 -----ThrLeuThrLeuSerAspLeuLysProSerA 963
Db 112124 CTTCCTCTCAAGGCTGCTTCTTCTCCTCAGACCCCTACATTTGACCGTGTGAGTCCCTCTC 112065
QY 963 rGlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlyS 983
Db 112064 GGGGCCAGCTTCTGGAGGCACCTGGCTTACCATCTCTGGAATTTCTCTGGATGCGGCA 112005
QY 983 erAsnValValMetPheGlyLysGlnProCysLeuPheHis----- 997
Db 112004 GCAGGGTCACAGTGAATTATAAGAGATGGCGAGTGGCCAGTTGT -GAGGTGGCGCTGAGCA 111946
QY 997 ----- 997
Db 111945 CAAGCCATCTGGAGTTGAGCACTGATTGTCTGAAGACTATGGGGCAGGAGGTGAAG 111886
QY 998 -----ArgArgSerProSerTyrI 1004
Db 111885 GTGCCAGGCCATCTGCTTTCATGCTTACCTGTGTAGGAGAGATGCCAGGCAA 111826
QY 1004 leValCys---AsnThrThrSerSerAspGluValLeuLysValSerValGlnV 1023
Db 111825 TCGTGTGTATCTACCTGCTCCACCTGGCCGCCAGTCAGAGCCCTATCACCTTGCCA 111766
QY 1023 alAspArgAlaLysIle---HisGlnAspLeuValPheGlnTyrValGluAspProThrI 1042
Db 111765 TCGATCATGCCAATCTCCAACTGAGTCACTATACCTACACCCAGGACCCCTACGG 111706
QY 1042 leValArgIleGluProGluTrpSerIleVal----- 1052
Db 111705 TCACACACCTTGAGCCACCTGGAGCATCATP -CAAGTGAGACCCCTGAGGAGATGGGGTT 111647
QY 1052 ----- 1052
Db 111646 GAGGACTGAGTTGGGGCTGTGAGAGTCTTGGGGCTTACCTATTACTGGGGCTTGGGTC 111587
QY 1052 ----- 1052
Db 111586 ATCAAGTGGTTCCTCTCTGTAGCCCTCATCAGTAGGGGCCACAGAGTTGTCTTGCCCA 111527
QY 1052 ----- 1052
Db 111526 TATAGAGCTGGCACCTGCTGAATGAACCAAGGTGAGAGGCTCTGAGCTCTCAGAAAATC 111467
QY 1053 -----Serg 1054
```

---

```
Db 111466 CTGTGTAATCTTGGACAGGGTGGGAAAGGAGTCTGAGCTCTTTACCCATCTCTGCACTG 111407
QY 1054 lyAsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleA 1074
Db 111406 GAAGCACCTCCATCACTCTGAGTGGNACCCATCTGCTGACAGTCTCAAGAACCCGCTGAC 111347
QY 1074 rGAlaLysHisGlyGlyLysGluHisIleAsnIle----- 1085
Db 111346 GAGCAAGTACCCTGCTATTAGACCACTAAGCTAGTAGTGGTCTTACCCCTCTGTTCT 111287
QY 1086 -----CysG 1087
Db 111286 ACCATTGCTACTCATGGCTTCATGGCTGATGGCTGAGATGCTGTGTCTTTCACAGACATGCC 111227
QY 1087 luValLeuAsnAlaThrGluMetThrCysGlnAlaProAlaLeuAlaLeuGlyProAspH 1107
Db 111226 AGGTGATCAACGACACACTGCAATGCTATGTAAAGCCCTTGGCATCTTCTTGGGACCCCTC 111167
QY 1107 isGlnSerAspLeuThrGluArgProGluGluPheGlyPheIleLeuAspAsnValGlnS 1127
Db 111166 AGCCTCGGGCCCCAAGCGGACGACCCCTGACAGTCTTGGCTTCTTGGTGGACCATGTCAGG 111107
QY 1127 erLeuLeuIleLeuAsnLysThrAsnPheThrTyrProAsnProValPheGluAlaP 1147
Db 111106 CAGCCCGCTCCCTCAACCCGTTCTTCTTCACTTACTTACCTGATCCAGCTTTGAACCCAC 111047
QY 1147 heGlyProSerGlyIleLeuGluLeuLysProGlyThrProIleIleLeuLys----- 1164
Db 111046 TTGGGCCCTCTGCTGTGCTAGATGTCAAACCTGGTTCACATGTTGTATTGAAGGTGCTGG 110987
QY 1164 ----- 1164
Db 110986 GGATAAGGATAGGAGTAGAAGAGGAGGGAACCCAGAGCTGGAGTGGTCTTGCATGTTT 110927
QY 1165 -----GlyLysAsnLeuIleProProValAlaGlyGlyAsnV 1177
Db 110926 CTGTGTGGTGTATTCACACTACTAGGGCAAGAACCTGATCCCTGCTGACGCTGGCAGCTCC- 110868
QY 1177 alLysLeuAsnTyrThrValLeuValGlyLysProCysThrValThrValIserAspV 1197
Db 110867 --CGCTCAACTACACAGTGTGATTGGAGGACGACCATGTGCACATACTCTACATA 110810
QY 1197 alGlnLeuLeuCysGluSerProAsnLeuIleGlyArgHisLysValMet----- 1213
Db 110809 CTCAACTTCTGTGTGATTCCTCCCAAGCCAGACAGCGCGGAGCTGTATTGTGGGTATGG 110750
QY 1213 ----- 1213
Db 110749 GCTCAAGAGGCCACCAGGCGCCCTGGCTGTCCATGGGAGTTGGTTCACCTCAAGCTTG 110690
QY 1214 -----AlaArg-ValGlyGlyMetGluTyrSerProGlyMetValTyrIleAlaPro 1230
Db 110689 CCTTACAGGTGCTGTGGTGGCTGGCTGAGTCTGTGGTGGCACGCTGCACATCACTGCT 110630
QY 1231 AspSerProLeuSerLeuProAlaIleValSerIleAlaValAlaGlyGlyLeuLeuIle 1250
Db 110629 GATGGGGCACTGACCTTACCAGCTATGGTGGGGCTAGCAGGAGGGGGGGCTTCTTCTG 110570
QY 1251 IlePheIleValAlaValLeuIleAlaTyrLysArgLysSerArgGluSerAspLeuThr 1270
Db 110569 CTGGCCATCACATGTTGTGTGTCGCTACACGCTAAGACTCAGGATGAGGACCGCAC 110510
QY 1271 LeuLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAlaLeuGluCysLys 1290
Db 110509 CTTAAGCGCTTTCAGCTACAAAATGGACAACCTGGAGTCTGCTGGCTTGGAGTGAAG 110450
QY 1291 Glu----- 1291
Db 110449 GAAGGTGCTAAAGCAGGGGGTGTACCATTTGTGTGAGAGCTGAGGAGCTTCTCTGCT 110390
QY 1292 -----AlaPheAlaGluLeuGlnThrAspIle 1300
```

```
Db 110389 CTTGTTCACTCACTACCGTATTGTTTCCCGCCCTTCTGCTGAGCTGCAGACTGATATC 110330
QY 1301 HisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspTyrArgThrTyr 1320
Db 110329 AATGAACACCAACACATGGATGGTGTCAGATCCCTTCTTGGACTACCGACCTAT 110270
QY 1321 ThrMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLeuGluVal 1340
Db 110269 GCTGTGGGTGCTTTTCCCTGGCATCGAGGCTCACCCAGTCTCAAGGAGCTGGATGTG 110210
QY 1341 -ProGlyTyrArgGlnGlu----- 1346
Db 110209 AGCCTCTACCCCAAGACCTGCCATATTCTGTTTCCCGACAGCCACCACTACCTACT 110150
QY 1347 -----ArgValGluLysGlyLeuLysLeuPh 1355
Db 110149 CTCTGAGATTCCCTGGTCCCCACAGACTCCCCCAATGTGGAGAGGCCCTTGGCTCTATT 110090
QY 1355 eAlaGlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGI 1375
Db 110089 TGGACAGTTGCTGCACAGCGCTGCTTCTGCTCACCTTATCCACACTTTGGAGGCCCA 110030
QY 1375 nArgSerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGI 1395
Db 110029 GAGTAGTTTCTCCATCGTGACCGTGGTACTGTGGCCCTCACTCACCATTGGTGGCCCTGCA 109970
QY 1395 nSerLysLeuGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLy 1415
Db 109969 GAGCCGGCTGACTATGCCACTGGCTGGCTGCTCAGCAACTGCTGGCTCACTCATAGAAA 109910
QY 1415 sAsnLeuGluSerLysAsnHisProLysLeuLeuLeuArg----- 1428
Db 109909 AAACCTTGAGAGCAAGAACACCCAAAGCTGCTATTGCGCAGGATGATGAAGTCTTGTCT 109850
QY 1428 ----- 1428
Db 109849 TGGCCCTCTCCTCAGTCTGCTTCCATTCTGCTTGGGCCCTCACTCTGTGTCAAGTC 109790
QY 1429 -----ArgThrGluSerValAl 1434
Db 109789 TCACCCTTGTGAGCCAGGSCCTTAGGACCTATCTCTTACCATAGGACAGACGCTGGTGGC 109730
QY 1434 aGluLysMetLeuThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLys----- 1451
Db 109729 TGAGAAGATGCTCACCACCTGGTTTCACATTCTGCTGATAGTTCTCGAAGTGTGTGCTG 109670
QY 1451 ----- 1451
Db 109669 GGGTGGCTGGGTGCGCAGGAGAGGAATGTCCGGGGGTGGGAGCAATTGATTGATAAGCCA 109610
QY 1452 -----GluCysAlaGlyGluProLeuPh 1459
Db 109609 GACAGGCAGCAGCTATGCCCTTAACCCCTGGGTAGGAGTGCGCCGGGAGCCACTCTT 109550
QY 1459 eSerLeuPheCysAlaIleLysGlnMetGluLysGlyProIleAspAlaIleThrGI 1479
Db 109549 CTTGCTGTACTGCGCCATCAAGCAGCAGATGGAAGAGGCTTATTGATGCCATAACAGG 109490
QY 1479 yGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnIleAspTyrLysTh 1499
Db 109489 CGAGGCCCGCTACTCTTAAGTGAGGACAAAGCTCATCCGGCAGCAGATCGACTATAAGAC 109430
QY 1499 rLeuVal----- 1501
Db 109429 CCTGGT-GAGCATGGGACTATGCGAGCAGGAAGTAAGTGGAGCCCTTGTAGGAATCAT 109371
QY 1502 -----LeuSerCysValSerProAspAsnAlaAsnSerPr 1513
Db 109370 GCTGAGGTTCTTCACTCTTAGACTCTCTGCATTGTGTGTCGCGAGAGCGAGGCGAGTGC 109311
QY 1513 oGluValProValLysIleLeuAsnCysAspThrIleThrClnValLysGluLysIleLe 1533
Db 109310 CCAGGTCCCTGTGAAGGTTCTTAACGTGTAGATCATCCCGCAGGCCCAAGACAGCTGTT 109251
```

```
QY 1685 ----- 1685
Db 108173 TAGACACCTTCTAGCCAGCTTTAAACATGAGATGAGTGTCTTCTTAGGCCCTTGGCAC 108114
QY 1685 ----- 1685
Db 108113 TAGGACCTCTACGCTGAGGTCAAGCAAGAGTCTATTATTTAGGAATGTTGGCCCTTCAGAG 108054
QY 1685 ----- 1685
Db 108053 AAGCTTTTTTTTACTGAGCTATTCTGCCCTCTGCTCAGATTTTGGAACTTCCTGAC 107994
QY 1686 ----- GlyThrLeuGlnLysPheValAspLeuPheG1 1697
Db 107993 ATGGTTCTTTTACCATCACTCATAGGGCAGCTTTCAGAAGTTTGTAGATGACTTGTGTTGA 107934
QY 1697 uThrIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPh 1717
Db 107933 AACTGTGTTCAGTACAGCCACCAGGGCTCAGCCTTACCCTTGGCCATCAAGTACATGTT 107874
QY 1717 eAspPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisTh 1737
Db 107873 TGACTTCTCGATGAACAGAGCTGACCAGCCAGATCAGTCAGCCGTGATGCGCTCACAC 107814
QY 1737 rTrpLysSerAsn----- 1741
Db 107813 CTGGAAGAGCAACTGTTACTGTTCTGACTGCTGCTGCTGCTTCTTCTGAGGGTGGGT 107754
QY 1741 ----- 1741
Db 107753 GGGAGGAGAGGCTAGTTACTTAGCCAAAGCCTCCAGCCCAATGGAGACAGGTTCACTAGC 107694
QY 1741 ----- 1741
Db 107693 AACTGAATAGCCTTCCATGCAGGAGGATGGACCAGTGGTATGSCCTGGAAGTGAATATG 107634
QY 1741 ----- 1741
Db 107633 AGTATATAGAAGTGGGTGCTGCTGGGTGTGTCAGATGGGAAGTGTAGTCTAGCACTGG 107574
QY 1741 ----- 1741
Db 107573 GATTGACCATTTCCTCTAGGGTCTGTCAGGGTGCAGGCTAGACACACTGATGGCTTGGCATAGG 107514
QY 1741 ----- 1741
Db 107513 TGGAGACAGTGGAAAGTTTCAGGGGTGAGGCTTGAGGACAGAAAGCCATCTATTCTTATTT 107454
QY 1741 ----- 1741
Db 107453 CTCAATTGCTGAGTGCAGCAGAGAGGTTAAAGAAACCCATTTATCCAAATAGTCTCTC 107394
QY 1741 ----- 1741
Db 107393 CTGTTACTTCAGGACAAGATATGACCTGGCCAAACAGGACAGCAAGGTGTGAAGGAGGCAGT 107334
QY 1742 ----- CysL 1743
Db 107333 GGGGGAAGGGGCACATAGGGCCATTAGGGAGTCTCAAGTGTATGTCCTCTGTTCAAGCT 107274
QY 1743 euProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheValPheAspIleHisL 1763
Db 107273 TACCTCTGCGTTTCTGGGTGAATGTGATCAAGAAATCCGCAATTTGTGTTTGACATCCATA 107214
QY 1763 ysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnThrPheMetAspSerCysS 1783
Db 107213 AGAATAGCATCACAGAGCCCTGTTGTGTCAGTGGTGGCCAGACCTTCATGGACTCCTGTT 107154
QY 1783 erThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysA 1803
Db 107153 CTACATCAGACACCGCGCTGGCAAGGACTCACCTTCCACAAAGCTCCTCTATGCCAAGG 107094
QY 1803 spIleProSerTyrLysAsnTrpValGlu----- 1812
```

```
Db 107093 ATATTCCCAATTACAAGAGCTGGGTGGAGAGGTGGCCCTGGGATACTGCTGTGTATGCC 107034
QY 1812 ----- 1812
Db 107033 TGGTGGCAGGAGGAAGTCCCTAAACAGCCTGTCTCAGGCTGGCCAGATACAAAGAAAAG 106974
QY 1812 ----- 1812
Db 106973 CCTCAGAGGCTGGCCAAACACTGATGGCATGAGTTTTTGGCATTTATCACCTGTGCATGAT 106914
QY 1813 ----- ArgTyrTyrSerAspIleGlyLysMetProAlaIleSer 1825
Db 106913 TCACCTACCCCAAGGCTCTCTATAGTACTATCGAGATATCGCAAAGATGGCATCCATCAGT 106854
QY 1826 AspGlnAspMetAsnAlaTyrLeuAlaGluGlnSerArgMetHisMetAsnGluPheAsn 1845
Db 106853 GACCAGGACATGGAGCGCTACTTAGTGGAGCAGTCCCGCTCCATGCTAATGACTTCAAT 106794
QY 1846 ThrMetSerAlaLeuSerGluIlePheSerTyrValGlyLysTyrSerGluGluIle 1864
Db 106793 GTCCTAAGTGCACCTCAGCGAGCTCTACTTCTATGTACCAAGTACCGTCAGGAGGTA 106737
```

Search completed: June 28, 2003, 15:21:45  
Job time : 13967 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 04:01:59 ; Search time 830 seconds

(without alignments)  
5144.324 Million cell updates/sec

Title: US-09-964-956-13

Perfect score: 9990

Sequence: 1 MKAMPWNWTCLLSHLLMVG.....QKLAYKLEQVITLMSLSDS 1896

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09964956/runat\_24062003\_124338\_5195/app\_query.fasta\_1.2055  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -NORM=score-pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09964956.cgn\_1.578 -runat\_24062003\_124338\_5195 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6794	68.0	6387	24	ABN59899
2	6765	67.7	6730	24	AB199504
3	3641	36.4	6382	23	ABL17265
4	2707	27.1	6408	22	AAS00024
5	2580	25.8	6252	22	AAH81745
6	2580	25.8	6252	22	AAS00020
7	2580	25.8	6252	22	ABL67279
8	2539	25.4	6329	22	ABA09062
9	2539	25.4	6329	22	AAH98400
10	2471	24.7	7080	22	ABA08953
11	2463.5	24.7	5892	22	AAS00022
12	2357.5	23.6	6561	23	AAS85204
13	1853.5	18.6	5367	22	AAS00021
14	1573.5	15.8	1088	22	AAH08997
15	1501	15.0	2971	22	AAK94342
16	1384	13.9	2414	22	AAK94356
17	1372	13.7	2597	24	ABK70006
18	1359.5	13.6	2477	20	AAK52244
19	1359.5	13.6	2477	21	AAC58585
20	1359.5	13.6	2477	21	AAA77558
21	1359.5	13.6	2477	22	AAS21409
22	1359.5	13.6	2477	22	AAF72402
23	1349	13.5	4707	20	AAK55560
24	1349	13.5	4707	21	AAC63721
25	1349	13.5	4707	22	AAS00320
26	1349	13.5	4707	22	AAS00459
27	1349	13.5	4707	22	AAF76945
28	1349	13.5	4707	22	AAF25249
29	1100	11.0	1844	24	ABQ54414
30	1095	11.0	601	22	AAS00023
31	973.5	9.7	995	23	ABV23162
32	973.5	9.7	995	23	ABV29002
33	925	9.3	598	22	AAK11122
34	925	9.3	598	22	AAK36939
35	925	9.3	598	22	AAI42745
36	920	9.2	2730	23	ABL18827
37	862	8.6	495	22	AAK23792
38	862	8.6	495	22	AAK43882
39	862	8.6	495	22	AAI58806
40	785.5	7.9	5357	23	ABL18826
41	779	7.8	932	24	ABN74259
42	681	6.8	2190	22	AAS00025
43	653	6.5	768	22	AAH03972
44	653	6.5	1767	22	AAH15398
45	629.5	6.3	4626	19	AAV20444

# ALIGNMENTS

## RESULT 1

ID ABN59899 standard; cDNA; 6387 BP.

XX ABN59899;

XX 28-JUN-2002 (first entry)

DT Novel human coding sequence SEQ ID NO: 310.

DE Human; antianaemic; vulnery; antiinflammatory; immunomodulator;  
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
KW expressed sequence tag; gene; ss.

XX Homo sapiens.

XX WO200222660-A2.

PN



1773	GAGACAGGTCACCAGGGTCCCGTGGAGTCATGTGGAGCAGTATACGACTTGTGGGGAGTG	1832	
QY	379	gGlyGluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisVallysValaGlyVa	899
Db	2910	TGGCGTGAACCTGGGTCTGGACTTCTCCGAGATGCCCCACCATGTGCAGGTGCGTGGGT	2969
QY	399	lGluCysSerProLeuValAspGlyTyrIleProAlaGluInIleValCysGluMetG1	919
Db	2970	GCCTTCAGCCCTCCAGGGGAATACATCATCGCTGAGCAGATTCTCTGTGAGATGGG	3029
QY	919	yGluAlaLysProSerGlnHisAlaGlyPheValGluIleCysValaAlaValCysArgPr	939
Db	3030	CCATGCCCTCGTGGGAACCACTCCGGGGCAGTACGCTGTGTATTGGCGAGTGTAAAGCC	3089
QY	939	oGluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLe	959
Db	3090	AGAGTTTCATGACGAAGTCCCATCAGCAGTACACCTTCGTGAACCCCTTCTGTGTGTCAC	3149
QY	959	uLysProSerArgGlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLe	979
Db	3150	CAACCCCAATCCGAGGTCCCGAGTCAGGAGCAGTATGGTGACCATTACCGGCCATTACCT	3209
QY	979	uAsnAlaGlySerAsnValValMetPheGlyLysGlnProCysLeuPheHisArgAr	999
Db	3210	TGGGGCTGGAGCAGCGTGGCAGTCTACCTGGGCAACCACTGCGAGTTCTACGGGAG	3269
QY	999	gSerProSerTyrIleValCys---AsnThrThrSerSerAspGluValLeuGluMetLy	1018
Db	3270	GTCAATGAGTGAGATCGTGTGTCTCACCCCATCATCCATGGCCTTGGCCCGGTCCC	3329
QY	1018	sValSerValGlnValAspArgAlaLysIleHisGlnAspLeuValPheGlnTyrValG1	1038
Db	3330	TGTTTCTGTGAGTGTGCGCCGAGCCCATGTGGATGACCACTGCAGTTTTCAGTACATAGA	3389
QY	1038	uAspProThrIleValArgIleGluProGluTyrSerIleValSerGlyAsnThrProI1	1058
Db	3390	TGACCTCGGGTCCAGGCATCGAGCAGAGTGGAGCATTCGACGTGGCCACACACCCCT	3449
QY	1058	eAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisG1	1078
Db	3450	GACCATCACAGGCTTCAACCTGGATGTCAATTCAGGAGCCCAAGGATCCGAGTCAAAATCAA	3509
QY	1078	yGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAl	1098
Db	3510	TGCAAGAAGATCTGTCATGTGTAAAGTTGTGAACACCAACCCCTCACCTGCCTGGC	3569
QY	1098	aProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPh	1118
Db	3570	ACCTCTCTGACCAGGACTACCGCCTGGACACTGTGGAGCCCGCAGATGAGTT	3629
QY	1118	eGlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsnPheThrTy	1138
Db	3630	TGGATTGTCTTAAACATGTCCCAATCCTTGCTTAATTTACACGACACCAAGTTTATCTA	3689
QY	1138	rTyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeuLysProG1	1158
Db	3690	CTACCCCAACCCGACCTTTCAACTGCTTAGCCCTACTGGAGTCTTGGATCAAAAGCCAGG	3749
QY	1158	yThrProIleIleLeuLysGlyLysAsnLeuIleProProValAlaGlyGlyAsnValLy	1178
Db	3750	ATGCCCATCATTTCTGAAGGGCAAAACCTCTGCCCTCTCGCTCTGGAGGG---GCCAA	3806
QY	1178	sLeuAsnTyrThrValLeuValGlyLysProCysThrValThrValSerAspValG1	1198
Db	3807	ACTCAACTACATGTCTCATCGAGAGACCCCTTGTGCTGTCTCAGCATCTATCTAGACCCA	3866
QY	1198	nLeuLeuCysGluSerProAsnLeuIleGlyArgHisLysValIleAlaArgValGlyG1	1218
Db	3867	GCTTCTCTCGAGCCCTCCCAACCTCACCGGGCAGCACAAAGTCTTGGTTCCAGTGGCGG	3926
QY	1218	yMetGluTyrSerProGlyMetValTyrIleAlaProAspSerProLeuSerLeuProAl	1238
Db	3927	GATGGTGTCTCCCTCGGCTCGGTGAGTGTCTCTCAGACAGCTTGTCTACCTGCGCAGC	3986

Qy	1238	alleValSerIleAlaValaGlyLeuLeuLeilePheIleValaValLeuI	1258
Db	3987	CATCGTCAGCATCGCGCGCGCGAGCTCCCTCCATCATCGTCATCATCGTCCTCAT	4046
Qy	1258	eAlaTyrlYsArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMe	1278
Db	4047	TGGCTACAAGCGGAAGTCTCGAANAATGACCTCTCAAGGGCTCGAATGCAGAT	4106
Qy	1278	tAspAsnLeuGluSerArgValAlaLeuGluCysLysGluAlaPheAlaGluLeuGlnTh	1298
Db	4107	GGCAATCTGGAGTCCGCTGGAGTCCAAGGAAGCTTTTGTGAGCTCCAGAC	4166
Qy	1298	rAsPileHisGluLeuThrSerAspLeuaspGlyAlaGlyIlepropHeLeuAspTyrAr	1318
Db	4167	GGATATCAATGAGTTGACCATGACCTGGACCGCTCAGGAATCCCTTACTTGCAGATCG	4226
Qy	1318	gThrTyThrMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLe	1338
Db	4227	TACCTAGCTATGGAGTCTGTTCCGGGCATCAGGACCACCCGCTCTCGGGGAGCT	4286
Qy	1338	uGluValProGlyTyrrArgGlnGluArgValGluLysGlyLeuLysLeuphealGlnLe	1358
Db	4287	GGAGGTACAAGAAACGGCAGCAGCAGTGGAGAAGCCCTCAAGCTCTTTGCCAGCT	4346
Qy	1358	uileAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPh	1378
Db	4347	CATCAACAAAGTGTTCTCTGTGACCTTCATCCGCACCCTCGAGCTCGAGCGCAGTTT	4406
Qy	1378	eSerMetArgAspArgGlyAsnValAlaSerLeulleMetThrValLeuGlnSerLysLe	1398
Db	4407	CTCCATCGCGACCGGGCAACGTGGCTTCGCTCATCAGCCGGCTCGAGCGCGCCT	4466
Qy	1398	uGluTyrrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleaspLysAsnLeuG	1418
Db	4467	GGATATGCCACTCATGTCTCAAGACGTGCTCTTGACCTCATCATTAAGAACCTTGA	4526
Qy	1418	uSerLysAsnHisProLysLeuLeuLeuArgThrGluSerValaGluLysMetLe	1438
Db	4527	GAACAAGAACCACCCCAAGCTGCTACTCCGGAGGACAGTCTGTGGCTCAAAAGATGCT	4586
Qy	1438	uThrAsnTrpPheThrPheLeuLeuTyrrLysPheLeuLysGluCysAlaGlyGluProLe	1458
Db	4587	GACCAATTGGTTCCGCTTCCTCCGCACAAGTTCTTAAGAGGTGCCAGGGAGGCACCT	4646
Qy	1458	uPheSerLeuPheCysAlaIleLysGlnGlnMetGluLysGlyProIleAspAlaIleTh	1478
Db	4647	CTTCATGCTATACTGTCCATCAAGCAGCAGATGGAGAAGGGCCCCCATTTGATGCCATCAC	4706
Qy	1478	rGlyGluAlaArgTyrrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLy	1498
Db	4707	GGCGAGCGCGCTACTCCCTGACGAGGACAAGCTCATCCGGCAGCAGATCCAGTACAA	4766
Qy	1498	sThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLy	1518
Db	4767	GACCTGTATCTGAACTCGCTCAACCCCTGACAACAGACAGTCCAGATCCAGTGAA	4826
Qy	1518	sIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuaspAlaIlePheLy	1538
Db	4827	GGTGTAAACTGTACACCATCACACAGTCAAGGAGAAGATTCCTGATCCCGTGTATAA	4886
Qy	1538	sAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTrpArgGlnG	1558
Db	4887	GAATGTGCCCTATTCCACGCGCGCGAGGCGAGTGGACATGGACTTGGAGTGGCCCAAGG	4946
Qy	1558	ySerGlyAlaArgMetIleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTr	1578
Db	4947	CCGGATCCCGGGTGTGCTGCAAGATGAGGACATCACCACCAAGATTTGAGGGTGACTG	5006
Qy	1578	plysArgLeuAsnThrLeuAlaHisTyrrGlnValProAspGlySerValValAlaLeuVa	1598
Db	5007	GAAGCGGCTCAACACCTGATGCATTTATCAGGTGTACAGCAGGTGGTGGTCTGGT	5066
Qy	1598	lSerLysGlnValThrAlaTyrrAsnAlaValAsnAsnSerThrValSerArgThrSerAl	1618

Db	5067	CCCCAACACAGCCTTCCTTACAAACATCCCTGCTCTGCAGCATCTCCCGGACGTCAT	5126
Qy	1618	aSerLysTyrGluAsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgTh	1638
Db	5127	CAGCAGATATGACTCCTCCTTACAGTATACGGCAGCCCGACAGCTCGGTCCTCCGGC	5186
Qy	1638	rPrometIleThrProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHi	1658
Db	5187	CCCGATATACCCACAGCTGGAAGTGGGGTCAAGGTGGGATCTGGTGAAGAACCA	5246
Qy	1658	sGluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLe	1678
Db	5247	TGACCACGGTGACCAGAAGGAGGTGACCGGGGACAGATGCTGTCCGAGATCTACCT	5306
Qy	1678	uThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspAspLeuPheGluTh	1698
Db	5307	GACCGGCTACTGGCCACCAAGGCAACCCCTGACAAAGTTTGTGACGACTGTGTTGAGAC	5366
Qy	1698	rIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAs	1718
Db	5367	CTTGTTGAGCACTGTGACCGGGGACGGCTCTCCCTGCCATCAAGTACATGTTTGA	5426
Qy	1718	pPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTr	1738
Db	5427	TTTCTAGATGAGCAGGACAGACAGCAGCATCCATGACACAGATGTGCGGCACACCTG	5486
Qy	1738	pLysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheVa	1758
Db	5487	GAAAGCAACTGCTCCCTCGCGCTTCTGGGTGAGCTGATTAAGAACCCCAAGTTTCGT	5548
Qy	1758	lPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnThrPh	1778
Db	5547	GTTTGACATCCACAAGGACACATCAGGACGCTGCTCTCTGTGTGGTGGCCAGACCTT	5606
Qy	1778	eMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLe	1798
Db	5607	CATGGACTCTGTTCAAGCTCAGAGCACCGGCTGGGCAAGGACTCCCCCTCCACAAGCT	5666
Qy	1798	uLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSerAspIl	1818
Db	5667	GCTCTATGCCAAGGACATCCCGCTACAGAGCTGGGTGGAGAGATACTACGACAGAT	5722
Qy	1818	eGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrIleuAlaGluInSerAr	1838
Db	5727	CGCCAAGCTCCAGCCATCAGTGACCGAGCATGAATGCCCTACCTCGCGAGCAGTCCCG	5786
Qy	1838	gMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValcl	1858
Db	5787	CCTGCACCCGCTGGAGTTCAACATGCTGAGTGCCCTCAATGAGATCTACTCCTATGTGAG	5844
Qy	1858	yLysTyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCysGlyLysGlnLy	1878
Db	5847	CAAGTATATAGTAGGAGGTCATCGGGGGCCCTAGACGAGATGAGCAGGACCGCGCAGCG	5906
Qy	1878	sLeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerLeuAspSer	1894
Db	5907	GCTGCTTATAGGTGGAGCAGCTCATTAATGCCATGTCCATGTGAGAGC	5955
RESULT 2			
ABI199504			
ID	ABI199504	standard; cDNA; 6730 BP.	
XX	AC	ABI199504;	
XX	DT		
XX	TX	07-MAR-2002 (first entry)	
XX	DE	Mouse ischaemic condition related cDNA sequence SEQ ID NO:479.	
XX	KW	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;	
XX	KW	vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.	
OS		Mus musculus.	

XX WO20018188-A2.  
 PN XX  
 PD XX  
 XX XX  
 PF 18-MAY-2001; 2001WO-JP04192.  
 XX XX  
 PR 18-MAY-2000; 2000JP-0145977.  
 XX XX  
 PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
 XX XX  
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
 XX WPI; 2002-034733/04.  
 DR P-PSDB; ABB57196.  
 DR XX  
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
 PT expression levels of particular genes defined in the specification or  
 PT by determining the expression profile of a gene group comprising these  
 PT genes -  
 XX  
 PS Claim 2; Page 1319-1331; 2690pp; English.  
 XX  
 CC The present invention describes a method for examining ischemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (1) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (1). The method  
 CC is useful for examining the ischemic condition (e.g. compressive  
 CC ischemia, occlusive ischemia or vasospastic ischemia) by measuring  
 CC expression levels of particular genes (ABI999202 to ABI99912, encoding  
 CC the protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The  
 CC expression levels or expression profiles produced by these genes are  
 CC used as an indicator when screening for ischemic condition-improving  
 CC drugs or therapeutics for ischemic diseases. ABI99913 and ABI99914  
 CC represent PCR primers for a mouse ischemic condition related sequence,  
 CC which are used in the exemplification of the present invention.  
 XX  
 SQ Sequence 6730 BP; 1656 A; 1850 C; 1740 G; 1482 T; 2 other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 6730  
 Score: 6765.00 Matches: 1268  
 Percent Similarity: 80.24% Conservative: 239  
 Best Local Similarity: 67.52% Mismatches: 357  
 Query Match: 67.72% Indels: 14  
 DB: 24 Gaps: 8

US-09-964-956-13 (1-1896) x ABI99504 (1-6730)

QY 25 LeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPheArg 44  
 DB 628 CTGCTAGCTCGTGGGACCACAGATGCTCCTCAG-----TACAGCACTTCCAC 675

QY 45 GlyGluProAlaGlu---GlyPheAsnHisLeuValValAspGluArgThrGlyHisIle 63  
 DB 676 TCTGAGAAATCGTGACTGGAGCTTTCACACCATTTGACTGTACACCGAAGACGGGCTGTG 735

QY 64 TyrLeuGlyValAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThr 83  
 DB 736 TATGTGGGGCTATCAATCTGCTGCTACAGTTGACTGGCAACCTCACCTCCAGGTGGCT 795

QY 84 HisGluThrGlyProAspGluAspAsnProLysCysTyrProArgIleValGlnThr 103  
 DB 796 CACAAGACAGGGCCAGAGAGGACACAAAGGCTTGTTACCCACCCCTCATTTGACACCC 855

QY 104 CysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIleAspTyrLys 123  
 DB 856 TGCAGTGAAGTGTCTACACTCACCAACAATGTCAACAATACTACTGATCATTTGACTACTCT 915

QY 124 GluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeuLeuArgLeu 143  
 DB 916 GAGAATCGGCTGTGGGCTCGCGAAGCCCTTACAGGGTGTGTTGCAAGCTCCTCGGACTA 975

QY 144 GluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeuSerGlyVal 163  
 DB 976 GATGACCTCTTCATCTGCTGGAGCCATCCACAGAAGAACACTACTTGTCCAGTGTC 1035

QY 164 AsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAspLysLeu 183  
 DB 1036 AATAAGACAGGCACCATGTATGCTGTGATGTGCGCTCTGAGGGGAAGATGCAAGCTT 1095

QY 184 PheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSerArgLys 203  
 DB 1096 TTTATCGGCACTGCTGTGGATGGCAAGCAGGATTACTTCCCTACTCTGTCCACGCCCAAG 1155

QY 204 LeuThrLysAsnSerGlyMetPheAlaTyrValPheHisAspGluPheVal 223  
 DB 1156 CTGCCCCGTGACCTGAGCTTTCAGCAATGCTGGACTATGACTCCACAGTATTGTGTC 1215

QY 224 AlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAspIleTyr 243  
 DB 1216 TCCTCCCTCATCAAGATTCCCTCTGACACCTAGCCCTGCTCTCTCACTTCGACATCTTC 1275

QY 244 TyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMet 263  
 DB 1276 TACATCTATGGCTTTGCCAGTGGGGGTTGTCTACTTCTCACTCTCCAGCCAGAG--- 1332

QY 264 ValSerProGly-----SerThrLysGluGlnValTyrThrSerLysLeu 280  
 DB 1333 ---ACCCCTGACGGCATGCCCATCAATTCAGCTGGAGACCTCTCTATACCTCAAGATT 1389

QY 281 ValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCys 300  
 DB 1390 GTGCGTCTCTGCAAGGATGACCCCAAGTTCCTACTGTGTGCTGCTTGTGGCTGTC 1449

QY 301 GluArgSerGlyValIleTyrArgLeuGlnAlaAlaTyrLeuSerLysAlaGlyAla 320  
 DB 1450 ACACGTGTGGGGTGGAAATATCGCTTCTGCAAGGACGTTACCTTCCAAAGCCAGGGGAA 1509

QY 321 ValLeuGlyArgThrLeuGlyValHisProAspAspLeuPheThrValPheSer 340  
 DB 1510 GCTCTAGCTAGCCCTTCAACATCAGCAGCGAAGATGCTCTGTGCTCCATCTTTTC 1569

QY 341 LysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeu 360  
 DB 1570 AAGGGCAGAGACGATACCCACCCCTGTATGACTGTGCTGCTGTGCTTCCCATC 1629

QY 361 LysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThr 380  
 DB 1630 CGGGCCATCACTTGCAAATCAAGGACGGTTGCGAGTCTCTTACCACGAGGAGGCAAC 1689

QY 381 LeuAspLeuAlaTrpLeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIle 400  
 DB 1690 TTGGAGCTCAACTGGCTGCTGGGAAGGATGTGCGAGTCGACCAAGCGCTGTCCCAATC 1749

QY 401 AspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArg 420  
 DB 1750 GATGATAAATCTTCGGCGCTGGACATCAACACCGCTCTGGGAGGCTCCACTCCTGTGGAG 1809

QY 421 GlyIleProValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrValTyr 440  
 DB 1810 GACTGACCCCTGTATACCAACAGCAGGACCGGCTGCTGTGCGCTCTCTATGTTTAC 1869

QY 441 LysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLysIleArgVal 460  
 DB 1870 AATGCTACAGTGTGTTTCTGGGACTAAGAGTGGCAAGCTGAAGAAGATTCAGCT 1929

QY 461 AspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGlnValVal---AspProGly 479  
 DB 1930 GATGTCCTCCCTCCATGAGTGGTCTCTGTGTGTCTGTTTCAAGATGGGAGC 1989

QY 480 ProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGlu 499  
 DB 1990 CCAATCTCCGGGACATGGCTTCTCCATCAATCACTATACCTATATGATGCTGTGAG 2049

Qy	500	ArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrGlnSerCysGlyGluCys	519
Db	2050	AGACAGTCCACGAGGTCCTGTTGAATCATGTGAACAGTATACAACCTTGTGAGAGTGT	2109
Qy	520	LeuGlySerGlyAspProHisCysGlyTyrCysValLeuHisAsnThrCysThrArgLys	539
Db	2110	CTAAGCTCAGGGATCCTCACTGTGGCTGGTGTGGCCCTGCACAACTGTGCTCCCGAAGA	2169
Qy	540	GluArgCysGluArgSerLysGluProArgArgPheAlaSerGlnMetLysGlnCysVal	559
Db	2170	GACAAATGCCACGGGCTCGGAAGCAAAATCGATTGCTGCCAGTATCAGCCAGTCGATG	2229
Qy	560	ArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuLeuValLeu	579
Db	2230	AGCCTTGAGGTACACCCCAACAGCATCTCTGTGTCAAGTACACAGCGCGGTGCTCAGCGCT	2289
Qy	580	GluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLeuSer	599
Db	2290	GTTGTGAATGATGCTCCCAACCTCTCTGAAGGTATTGCTTGTGCCCTTGGGAATCGACT	2349
Qy	600	GluMetAspGlyLeuValValGlyAsnGlnIleGlnCysTyrSerProAlaAlaLysGlu	619
Db	2350	GAGGTGAGGACAGGTATCTGGAGTCAGTCATCTGCATCTCACCTGGACCCAGGAT	2409
Qy	620	ValProArgIleIleThrGluAsnGlyAspHisHisValValGlnLeuGlnLeuLysSer	639
Db	2410	GTCCCT---GTCATCCCTCTGGATCAAGACTGGTTTGGCCTAGAGCTGCAGCTGAGATCC	2466
Qy	640	LysGluThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerValHis	659
Db	2467	AAAGAGACAGGAAGATCTTTTGTGCAGCACCGGAATTCAAGTTCATATACTGCATGGCCAC	2526
Qy	660	AsnSerCysLeuSerCysValGluSerProTyrArgCysHisTyrCysLysTyrArgHis	679
Db	2527	CAACTGTGCGCTGCTCTGTGTAAACAGCGCCTTCGCTGCCATTGGTGCAAGTACCGTAAC	2586
Qy	680	ValCysThrHisAspProLysThrCysSerPheGlnGluGlyArgValLysLeuProGlu	699
Db	2587	CTCTGCACACATGACCCCACTACCTGTCTCTCCAGGAAGCAGGATCAATGTTTCAGAG	2646
Qy	700	AspCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysPro	719
Db	2647	GACTGTCCCAAGCTGTGCGCCACGAGGAGATCTGTATCCAGTGGGGAAGTAAACCA	2706
Qy	720	IleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCys	739
Db	2707	ATCACCTTAAGGCCCAACCTGCCCCAGCTCGCCAGCAGGCTACGAGTGT	2766
Qy	740	IleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArgPheAsnSerSerSer	759
Db	2767	GTGCTCAGCATTCGAAGGGGTGTCCACCGGGTCCCTGCCCTGGGTTTCAACAGTTCAGT	2826
Qy	760	ValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsnLeuProVal	779
Db	2827	GTGCAGTGCACAAACAGCTCGTACCATGTATGATGGATGGACATCAGCAACCTCAGCAGTG	2886
Qy	780	GluLeuThrValValTyrAsnGlyHisPheAsnIleAspAsnProAlaGlnAsnLysVal	799
Db	2887	GACTTTGCTGTAGTATGATGGCAACTTCATTATTGACAACTCAGGACCTGAAAGTA	2946
Qy	800	HisLeuTyrLysCysGlyAlaMetArgGluSerCysGlyLeuCysLeuLysAlaAspPro	819
Db	2947	CATCTCTACAAGTGTGCAGCCACGCGGAAGCTGTGGTCTCTGCCCTCAAGGCTGACCA	3006
Qy	820	AspPheAlaCysGlyTyrCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysPro	839
Db	3007	AAGTTGAGGTGTGGCTGGTGCAGTGTGACGCGCAGATGTACCTCCACAGCACTGCCCC	3066
Qy	840	AlaGlnGlnSerGlnTyrLeuGluLeuSerGlyAlaLysSerLysCysThrAsnProArg	859
Db	3067	AGCACTTCTAGCCCTGGCTGACTGGTCCAGCCACAATGTCAGTGTTCACACCCCA	3126
Qy	860	IleThrGluIleIleProValThrGlyProArgGluGlyGlyThrLysValThrIleArg	879

Db	3127	ATCACAGATTTGACAGTATCAGGACCACTTCAGAGGAGGACTCGTGTGACCATCCAT	3186
Qy	880	GlyGluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyVal	899
Db	3187	GGCGTGAACCTGGCGTTGGACTTCTCTGAGATGCTCACCATGTGCAAGGTGGGTGGAGTG	3246
Qy	900	GluCysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMetGly	919
Db	3247	CCCTGACACATCCAGCGGGAATACTCATCGCTGACGATGCTGCTGTGAGATGGC	3306
Qy	920	GluAlaLysProSerGlnHisAlaGlyPheValIleCysValAlaValCysArgPro	939
Db	3307	CATGCCGTTATAGTACCACTCGGCCCTGCGCCTGTGCATTTGGTGGGAATGCAAGCCA	3366
Qy	940	GluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeu	959
Db	3367	GAGTTCATGACCAAGTCCCAAGCAGCATATCTTTGTGAATCCTCTGTGCTGTCACTC	3426
Qy	960	LysProSerArgGlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeu	979
Db	3427	AGCCCGATCCGGGACCAAGTACAGAGGTACCATGCTGACCATCAGGCCATTACCTT	3486
Qy	980	AsnAlaGlySerAsnValValMetPheClyLysGlnProCysLeuPheHisArgArg	999
Db	3487	GGTCTCGGAGCAGTGTGGCAGTCTACTCGGCAATCAGACCTGTGAATTTCTATGGGAGG	3546
Qy	1000	SerProSerTyrIleValCys--AsnThrThrSerSerAspGluValLeuGluMetLys	1018
Db	3547	TCAATGAATGAGATTGTATGTTTACCCCCATCATCCATGGACTAGGACAGTCCCT	3606
Qy	1019	ValSerValGlnValAspArgAlaLysIleHisGlnAspLeuValPheGlnTyrValGlu	1038
Db	3607	GTCTCCGTGAGTGTGCACAGACGCCGGGTGGATAGCAGCTGTGAGTTCGAGTATATAGAT	3666
Qy	1039	AspProThrIleValArgIleGluProGluThrPsrIleValSerGlyAsnThrProIle	1058
Db	3667	GACCCAGGGTCCAAAGCTATTGACCCAGAGTGGAGTATCACTAGTGGGCACACACCCCTA	3726
Qy	1059	AlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGly	1079
Db	3727	ACCATCACAGCGTTCAACTTGGATGTTCATTCAGGAGGCCAGGGTCCGAGTCAAAATTTAAT	3786
Qy	1079	GlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAla	1098
Db	3787	GGCAAGAAGTCTGCAATGTATGCACAGTGGTAAACACAAACCCCTCCTGCTGGCA	3846
Qy	1099	ProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPhe	1118
Db	3847	CCCTCTGTGACCATGACTACCGCCAGGTCTGGACACTGTGGAAACGCCAGATGAGTTT	3906
Qy	1119	GlyPheLeuLeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsnPheThrTyr	1138
Db	3907	GGATTTCTTTTAAACAATGTTTCAATCTTACTCATCTATACAGACACCAAGTTTCATCTAC	3966
Qy	1139	TyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeuLysProGly	1158
Db	3967	TACCCCAACCCACAGTTTGAACCTGCTCAGCCCCACTGGAACTTGGATTCAGAAGCCAGGC	4026
Qy	1159	ThrProIleLeuLysGlyLysAsnLeuIleProValAlaGlyGlyAsnValLys	1178
Db	4027	TCACCATCATCCTGAAGGGCAAAAATCTCTGTCTCTCTGCTCTGGAGG---GCCAAA	4083
Qy	1179	LeuAsnTyrThrValLeuValGlyGluLysProCysThrValThrValSerAspValGln	1198
Db	4084	CTCAACTACACAGTAATGATGGAGAGACACTGTATACAGTCACTGTGCTGACACACAC	4143
Qy	1199	LeuLeuCysGluSerProAsnLeuIleGlyArgHisLysValMetAlaArgValIcylGly	1218
Db	4144	CTGCTTTGTGAACCTCCCAACCTCACAGGGCAGCACAGGTCATGGTTCACGTGGCGGG	4203
Qy	1219	MetGluTyrSerProGlyMetValTyrIleAlaProAspSerProLeuSerLeuProAla	1238

Db 4204 ATGGTGTTCTACCTGGCTCCGTCAGCGTCATCTCCGACAGCCTGTTGACCCCTCCAGCC 4263  
QY IleValSerIleAlaValAlaGlyGlyLeuLeuIleIlePheIleValAlaValIle 1258  
Db 4264 ATCATCAGCATCGCAGCTGGTGAAGCCTCTTATCATCATCATCTGCTCATC 4323  
QY AlaTyrLysArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMet 1278  
Db 4324 GCTTACAAGCGCAAGTCTAGGAAAATGACCTCACACTCAAGCGCTCCAAATGCAAAATG 4383  
QY AspAsnLeuIleSerArgValAlaLeuGluCysLysGluAlaPheAlaGluLeuGlnThr 1298  
Db 4384 GACAACCTGGAGTCCAGGTGGCAGCTGGAGTGAAGAACTTTGCGGAGCTTCAGACA 4443  
QY AspileHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspTyrArg 1318  
Db 4444 GACATCAATGAGCTAACCACTGACTTGGATCGATCAGGAATCCCTTACCTGGACTACCGT 4503  
QY ThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLeu 1338  
Db 4504 ACCATATGCCATGAGAGTCTGTTCCAGGCATTCAGGACCACCTGTTCTGCGGGAAGCTG 4563  
QY GluValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeu 1358  
Db 4564 GAGGTACAGGGAATGGACAGCAGCAGCTGGAGAAAGCCCTGAAACTCTTCGCCCAAGCTT 4623  
QY IleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPhe 1378  
Db 4624 ATCAACAACAAGGTGTTCTGTGTCACCTTCATCCGTACACTGGAACCTACAGCGCAGCTTC 4683  
QY SerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeu 1398  
Db 4684 TCCATCGCAGACCGTGGGAAAGCTGGGCTCTCTCATCATGACAGCGCTTCAGGGTCGCCTA 4743  
QY GluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGlu 1418  
Db 4744 GAATATGCCATGATGCTCTCAAGCAGCTGCTCTGACCTCATTTGACAAGACCTCGAG 4803  
QY SerLysAsnHisProLysLeuLeuArgArgThrGluSerValAlaGluLysMetLeu 1438  
Db 4804 AACAACAACCCCAAGCTGCTTCCCGCAGGACTGAGTCTGTGGCCGAGAGAGTCTG 4863  
QY ThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeu 1458  
Db 4864 ACTAATGTTGTTCTTCTTACACAAGTTCCTGAAGAGTGTGCTGGGGAACCACTC 4923  
QY PheSerLeuPheCysAlaIleLysGlnGlnMetGluLysGlyProIleAspAlaIleThr 1478  
Db 4924 TTCATGCTATCTGTCATCAGCAGCAGATGGAAGAAAGGCCCTTATGACGCTATTACT 4983  
QY GlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLys 1498  
Db 4984 GGTGAGGCCGCGATACCTCCCTGAGTGAAGACAAGCTCATCCGCGCAGCAGATCGAGTAAAG 5043  
QY ThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLys 1518  
Db 5044 ACTCTGATCCTGAACCTGTGTCAACCTGCAATGAGAACAGCCAGCAGATCCCAAGTGA 5103  
QY IleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLys 1538  
Db 5104 GTACTAACTGTGACACCATCATCAAGTCAAGAGAGATCCTCGATGCCGTATATAAG 5163  
QY AsnValProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTyrArgGlnGly 1558  
Db 5164 AATGTCCCTTACTTCCCGAGCGGCAAGGCTGTGGACATGATCTGGAGTGGCGCAAGGC 5223  
QY SerGlyAlaArgMetIleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTrp 1578  
Db 5224 CGGATTTCCAGAGTGTGTGAGGACGAAGACATTACCACCAAAATAGAGGGTGAAGTGG 5283  
QY LysArgLeuAsnThrLeuAlaHisTyrGlnValProAspGlySerValValAlaLeuVal 1598  
Db 5284 AAGCGGCTTAACACTGATGCAATACCAGGTGTACAGAGATCCCGTGGTCTGCTGTT 5343

QY SerLysGlnValThrAlaTyrAsnAlaValAlaAsnAsnSerThrValSerArgThrSerAla 1618  
Db 5344 CCTAAGCAGACCTCTCTCTACCAACATCCCTGCCTCTGCCAGCATCTCTCGGACATCCATT 5403  
QY SerLysTyrGluAsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgThr 1638  
Db 5404 AGCAGATGATGACTCTCTTCCAGGTACACAGCAGCAGCCAGACAGCCTCCGGTCCCGGTC 5463  
QY ProMetIleThrProAspLeuSerGlyValLysMetTrpHisLeuValLysAsnHis 1658  
Db 5464 CCCATGATCACCCAGACTTGAGAGCGGTGTCAGGTTTGGCATCTGCTGGAAGAATCAT 5523  
QY GluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLeu 1678  
Db 5524 GACCATGTTGACCAAGAGGAGGTGACCGGCGCAGCAAAATGTTGTGACGACTTGTTTGAGACC 5583  
QY ThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspAspLeuPheGluThr 1698  
Db 5584 ACCCGGCTTCTAGCCACCAAGGACCTGCAAGAAATTTGTGACGACTTGTTTGAGACC 5643  
QY IlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAsp 1718  
Db 5644 TTGTTACACACTGTCCACCGGTAGTCTCTCCCTAGCCATCAAGTACATGTTGAT 5703  
QY PheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrp 1738  
Db 5704 TTCTGTGATGAGCAGCAGCAGACACAGTATCCAGCAGACAGATGTCGCGCACACCTGG 5763  
QY LysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheVal 1758  
Db 5764 AAAAACAATGCTTCCACTTCTGTTCTGGGTGAATGTCTCATCAAGAACCTCAATTTGTA 5823  
QY PheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnThrPhe 1778  
Db 5824 TTTGACATCCACAAGGCGCAGCATCAGATGCTGCTGCTGTGGTAGCCAGACCTTT 5883  
QY MetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeu 1798  
Db 5884 ATGGACTCTCTTCCACATCAGACGACCTAGGCAAGGACTCACCTTTCCAACAAGCTG 5943  
QY LeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrThrSerAspIle 1818  
Db 5944 CTCATGCAAGGATATCCCAAGTATATAAGAACTGGTAGAAGATACTATGCAGATATT 6003  
QY GlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGlnSerArg 1838  
Db 6004 GCCAAGCTCCAGCCTATTAGTACCAGATATGATGCTTACTCCTCGGAGCAGTCCCGC 6063  
QY MethHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGly 1858  
Db 6064 CTGCATGCTACAGAGTTCAATATGCTAGCGCCCTCAACGAGATCTACTCATATGTCAGC 6123  
QY LysTyrSerGluGluIleLeuGlyProLeuAspHisAspAspGlnCysGlyLysGlnLys 1878  
Db 6124 AAGTACATGAGGAGCTCATCGGCGCATAGAGGAGTGAACAGGCGCCGAGCGCAACGA 6183  
QY LeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerLeuAspSerAsnLys 1896  
Db 6184 CTGGCCTACAGGTGAGCATCTCATCAAGCCCATGCTCCATAGAGAGCTCAAG 6237  
RESULT 3  
ABLI17265  
ID ABLI17265 standard; DNA; 6382 BP.  
..XX  
AC ABLI17265;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 3268.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical; gene; ds.

XX Drosophila melanogaster.  
OS W0200171042-A2.  
PN XX  
XX 27-SEP-2001.  
PD XX  
XX 23-MAR-2001; 2001WO-US09231.  
PF XX  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX (PEKE ) PE CORP NY.  
PA Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
PT Claim 1; SEQ ID NO 3268; 21pp + Sequence Listing; English.  
PS The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABU16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 6382 BP; 2073 A; 1195 C; 1282 G; 1832 T; 0 other;

Alignment Scores:  
Pred. No.: 2,47e-290 Length: 6382  
Score: 3641.00 Matches: 829  
Percent Similarity: 58.15% Conservative: 320  
Best Local Similarity: 41.95% Mismatches: 637  
Query Match: 36.45% Indels: 190  
DB: 23 Gaps: 60

US-09-964-956-13 (1-1896) x ABL17265 (1-6382)

QY 26 LeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPheArgGly 45  
Db 280 TTGACTAATGCAATGCACCGGATTAAGAATGCAAGAATCTGAATTTCTACGATTACAAAT 339  
QY 46 GluProAla-----GluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIle 63  
Db 340 GTAGCTCGCTTGACACCAATTTGAATCACCATTATAGTAGGACACCATTTACTGGCAGGTC 399  
QY 64 TyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThr 83  
Db 400 TTCGTTGGTGGTTAACAGGTTGATCAGTTGCGCTGATTTGGAGCTTTCGGAACAT 459  
QY 84 HisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIleValGlnThr 103  
Db 460 GTGAAGACGGGGCTCAAAATGATTTTCGTCGAGTGT-----AGCATCTTGTAT 507  
QY 104 Cys-----AsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIleAsp 121  
Db 508.TGTCATTAAATGAGTTCGAGTCCGAGTCCGCGGACCAATATATAAGGCTTCTGCTCATTGAT 567  
QY 122 TyrLysGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeuLeu 141  
Db 568 CGTGCACACTTCAAGACTATTGCGTGGATGATCATTATTTCAAGGTACATGCACAGTTCTGT 627  
QY 142 ArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeuSer 161

Db 628 AATCTTCAAAATGTCAGCATAATT-----GAGCATGAAGTTCTCT 666  
QY 162 GlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerSerLeuAspAsp 181  
Db 667 GAT-----GCTGTTGGCTTAATGATGCCAACTCCTCAACC 702  
QY 182 LysLeuPheIleAlaThrAlaValAspGlyLysProGluTyr----- 195  
Db 703 GTAGCATTTATTGCG-----CCGGGACCTCCACAACATCCGGTGACTAATGTGATG 753  
QY 196 -----PheProThrIleSer 200  
Db 754 TACGTCGGAGTTACATACACAATAATTACCGTCACCGAGTGAATTCGCCGAGTTGCA 813  
QY 201 SerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAsp 220  
Db 814 TCCCGCTCCTCGAAGACTAAA-----ATGTTT----- 843  
QY 221 GluPheValAlaSerMetIleLysIleProSerAspThrPheThr-----IleIle 237  
Db 844 CAGATTGCTCTCGCGCAGTTACAACTGGAACGAGAACTTTTATAAACTCATACGCACGT 903  
QY 238 ProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeu 257  
Db 904 GAAACATATTTGTCAACTATCTTTATGATTCAGTTCTCGAGCATTTTCTTACTTTCTA 963  
QY 258 ThrLeuGlnProGluMetValSerProProGlySerThrThrLysGlyGlnValTyrThr 277  
Db 964 ACAACACAG-----TTAAGCACAGTCACCACTTCTTCTCTAAAGAG-----TACATA 1011  
QY 278 SerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValPro 297  
Db 1012 ACGAAACTCGTACGAATATGTCAGGAAGACTCGAAGTATTATTCTTACACTGAAATTCGA 1071  
QY 298 IleGlyCys-----GluArgSerGlyValGluTyrArgLeuLeuGlnAlaIaTyr 314  
Db 1072 GTGGAATGTATAAGTGATGCTCAAGCGCGCACTAAATCAACTTGGTTCCAGTGGGATTC 1131  
QY 315 LeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspLeu 334  
Db 1132 TTGGGAAACCTAGCTCAGACCTGGCTCAAAGCTTAGGTATATCTATTCAAGATGATGTC 1191  
QY 335 LeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSerAla 354  
Db 1192 CTCCTTGTGCTTTTTCAAAAGGAGAGGAAT-----ACACCAACGACAATTTCTGCA 1245  
QY 355 LeuCysIlePheIleLeuLysGlnLysAsnAspArgIleLysGluArgLeuGlnSerCys 374  
Db 1246 CTCCTCATTTACTCAATTAATCTATTCTGCGAAGCTTTATGCAACATATAAATCTCTGT 1305  
QY 375 TyrArgGlyGluGlyThrLeuAspLeuAlaTTPLeuLysValLysAspIleProCys--- 393  
Db 1306 TTTAACGGAAGTGGCATGAGGCGCTGGACTTTATATCA---CCCAGCATGCCATGTGTT 1362  
QY 394 SerSerAlaLeuLeuThrIleAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413  
Db 1363 CTAAAGAACTCAACACCATTTGGAGAGACTTCTGTGGATTAGATATAAATCACCCTTA 1422  
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433  
Db 1423 GGAGAGAAACCCCATTTACTTCAGTACCGGTG---GCCATGTTTACACAAAGTTAACT 1479  
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
Db 1480 TCTGTTGCTGCTACGAGTACAAGCGGTATACCGTGTCTTCTGCGGAGTACGCGATGGA 1539  
QY 454 LysLeuLysLysIleArgValAspGlyProArgGlyAsnAlaLeuGlnTyrGluThrVal 473  
Db 1540 TTTCTTAAGAAAGTTGTAATCGAGTCT---TCATCTATTGCCAACGAGTAGTAGTTT 1596  
QY 474 GlnValValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisGln 493

Qy	824	GlyTrpCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysProAlaGln-----	841
Db	2689	GGATGGTCTCGTCTACTAATACATGCGAAGTTCGAACAATTAATAAAAAACAAGAG	2748
Qy	842	--GluSerGlnTrpLeuGluLeuSerClyAlaLysSerCysCysThrAsnProArgIle	860
Db	2749	GGGAANAAGTATGGTTAAATCGAAGT-----GAAATTTGCCCAACCCCTGAAATT	2799
Qy	861	ThrGluTlleProValThrGlyProArgGluGlyGlyThrLysValThrIleArgGly	880
Db	2800	CATACATTCCAACCAACCAACAGGTCCTGGGAAGAGAGAACAAATAACAATAACGAGGA	2859
Qy	881	GluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyValGlu	900
Db	2860	ATTAATTTAGGCAAAATTAACAGATATTTATTCAGGAGTTGCAATAGCTGGTATAAT	2919
Qy	901	CysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMet-----	918
Db	2920	TGTATGCGGTTCTCTCAATTTATTTGACACTAAGCAAAATAGTATGTACCGTGGATAGT	2979
Qy	919	--GlyGluAlaLysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCys	937
Db	2980	CCTGGAGAACAGATGTATAGAAATGGCAAAATACGTGTCACAAATT-----	3024
Qy	938	ArgProGluPheMetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSer	957
Db	3025	--GGGACTACCGTGGTGATGTCGAAGAGGACTACGAGTTGCTGTACCCGCAAAATTTG	3081
Qy	958	AspLeuLysProSerArgGlyProMetSerClyGlyThrGlnValThrIleThrGlyThr	977
Db	3082	GATTTTAAATCCAAAGTTTGTGCCAACTTCAGTGGAACCGCAATATACATATACCTGGAAG	3141
Qy	978	AsnLeuAsnAlaGlySerAsnValValMetPheGlyLysGln---ProCysLeuPhe	996
Db	3142	CACTTAAATGCGGCTCTCGTATATACAGCATCCATAAAGCATCACTTACCGTGTAAAAATA	3201
Qy	997	HisArgSerProSerTyrIleValCysAsnThrThrSerSerAspGluValLeuGlu	1016
Db	3202	CTAAGCACAGATTCCTCAAGCAATATGTGCGACATTCCTGCTCCCTGGTATTTATTTGAA	3261
Qy	1017	MetLysValSerValGlnValAspAlaGala---LysIleHisGlnAspLeuValPheGln	1035
Db	3262	GGAAGACTAAAATGTCTTTTGATAATGGGCCCGAGAGTTTATGACTACAATTTTAAA	3321
Qy	1036	TyrValGluAspProThrIleVal-----ArgIleGluPro	1047
Db	3322	TATGTACTGGATCTACAGTTTGAACAGCTTAGTTCTGGCAACGGCGCAATAAAGATT	3381
Qy	1048	GluTrpSerIleValSerGlyAsnThrProIleAlaValTrpGlyThrHisLeuAspLeu	1067
Db	3382	CCAAAAGAAATACCAAGCAGCGGTATACGAATTTCTGTAAAGGTACCCCAATTTACCAGT	3441
Qy	1068	IleGlnAsnProGlnIleArgAlaLysHisGlyLysGluHisIleAsnIleCysGlu	1087
Db	3442	ATACAAATCCCAATATTATGTAGTGTATAATGGAGAGATGTATGCAAGCCCATGTCGG	3501
Qy	1088	ValLeuAsnAlaThrGluMetThrCysGlnAlaProAlaLeuAlaLeuGlyProAspHis	1107
Db	3502	GTGCAATCGGATACGGAATGGAATGTGCATCTCCAGTTGTTGATGTG-----GACAGC	3555
Qy	1108	GlnSerAspLeuThrGluArgPro-----GluGluPheGlyPheIleLeuAspAsn---	1124
Db	3556	CATGTTTATGAAGCGAAAGGCCAAATTTCTTGAATATGGTTCTCTTAATTTGGACAATGTT	3615
Qy	1125	-----ValGlnSerLeuLeuLeuAsnLysThrAsnPheThrTyrTyrProAsnPro	1142
Db	3616	TTGCGTGTTCAAAATTTATCTAAGTTTCATAACACCATTTCGAGCTTTACCCAAATCCA	3675
Qy	1143	ValPheGluAlaPheGlyProSerGlyIleLeuLeuLeuLysProGlyThrProIleIle	1162
Db	3676	GAGTATTTTATATTCGAAGAAGA-----GTTAAGTACTTTAAGAGCTGAATACTTAACC	3729



Db 5860 TACGCTGTGAACATGAACAACTAACTGTTACACTTGAAGAGGATGAATTTTCGCAA 5919  
 QY 1876 LysGlnLysLeuAlaTyrLysLeuGlnValIleThrLeuMetSer 1891  
 Db 5920 AAGCAGGATGGCTTTAAATTAGAGCAAGTTCACAACATAATGTCG 5967

## RESULT 4

AA500024  
 ID AA500024 standard; cDNA; 6408 BP.

AC AA500024;

DT 09-MAY-2001 (first entry)

DE Human cDNA encoding a Plexin.

XX Human; plexin; semaphorin domain; hyperplasia; neoplasia; cancer;  
 KW neurodegenerative disease; autoimmune disease; lupus; multiple sclerosis;  
 KW inflammatory bowel disease; diabetes type I; rheumatoid arthritis;  
 KW immunogen; antibody; ss.

XX Homo sapiens.

EH Location/Qualifiers

FT 1..6408

FT /\*tag= a

FT /transl\_except= (pos:3889..3891,aa:Ser)

FT /transl\_except= (pos:4873..4875,aa:Ser)

FT /product= "Plexin"

XX WO200114420-A2.

XX 01-MAR-2001.

XX 25-AUG-2000; 2000WO-US23365.

XX 25-AUG-1999; 99US-0150576.

XX (UYTO-) UNIV TORINO.

XX (REGC) UNIV CALIFORNIA.

XX Artigiani S, Comoglio PM, Goodman CS, Tesier-Lavigne M;

XX Tamagnone L;

XX WPI; 2001-226610/23.

XX P-PSDB; AAU00019.

XX New plexin polynucleotides and polypeptides, useful in diagnosis,  
 PT therapy and in producing compounds for treating diseases involving  
 PT aberrant cell growth (e.g. cancer) or immune regulation (e.g.  
 PT autoimmune diseases)

XX Disclosure; Page 70-72; 79pp; English.

XX The sequence encodes a Human Plexin. Plexins are large transmembrane  
 CC proteins whose extracellular domain shares homology with Scatter factor  
 CC receptors and contain an approximately 500 amino acid Semaphorin domain.  
 CC The plexin polynucleotides and polypeptides, and plexin-specific binding  
 CC agents are useful in diagnosis, therapy and in the biopharmaceutical  
 CC industry. In particular, the plexin polynucleotides and polypeptides are  
 CC useful for generating compounds (e.g. plexin-specific binding agents or  
 CC antibodies) for treating or diagnosing a disease or disorder involving  
 CC aberrant cell growth (e.g. hyperplasia, neoplasia, cancer or  
 CC neurodegenerative disease), or diseases or disorders involving aberrant  
 CC immune regulation (e.g. autoimmune diseases such as lupus, inflammatory  
 CC bowel disease or Diabetes Type I), or immunosuppressive diseases such as  
 CC multiple sclerosis or rheumatoid arthritis.

XX Sequence 6408 BP; 1190 A; 2015 C; 1948 G; 1255 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 4,66e-213 Length: 6408

XX Score: 2707.00 Matches: 689

Percent Similarity: 46.31% Conservative: 354  
 Best Local Similarity: 30.60% Mismatches: 701  
 Query Match: 27.10% Indels: 508  
 DB: 22 Gaps: 66

US-09-964-956-13 (1-1896) x AAS00024 (1-6408)

QY 12 LeuSerHisLeuLeuMetValGlyMetGlySerSerThrLeuLeuThrArgGlnProAla 31  
 Db 10 CTGGGCCAGCTCTTCTCCAGGCTCTCTGGGCCGGGTGGTCTTCCACCTCCAGCCCTT 69  
 QY 32 ProLeuSerGlnLysGlnArgSerPheVal-----ThrPheArgGlyGluProAla 48  
 Db 70 CCA-----CCAACGTGCATTCTACTCCCAATGCGAGTAT----- 102  
 QY 49 GluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIleTyrLeuGlyAlaVal 68  
 Db 103 -----CTGCAGCACCTGGCAAGGACCCCACTCAGGCACCTCTACTTGGGGCTACC 156  
 QY 69 AsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThrHisGluThrGlyPro 88  
 Db 157 AACTTCTCTTCCAGCTGAGCCCTGGGCTGCAGCTGGAGGCCACAGTGTCCACCGCCCT 216  
 QY 89 AspGluAspAsnProLysCysTyrProProArgIleValGlnThrCysAsnGluProLeu 108  
 Db 217 GTGCTAGACAGCAGGAGCTGCTGCCACCTGTGATGCTGTAGTGCCTCCAG-----GCC 273  
 QY 109 ThrThrThrAsnAsnValAsnLysMetLeuLeuIleAspTyrLysGluAsnArgLeuIle 128  
 Db 274 CAGCTTACCAACAACCCGAATCAGTGTCTCTCTGGTG-----AGCCCAAGGGCCCTGGTG 327  
 QY 129 AlaCysGlySerLeuTyrGlnGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 148  
 Db 328 GTATGCGGGAGCGTGCACAGGGGTCTGTGAACAGCGCGCTCGGGCAGCTCCAGCAG 387  
 QY 149 Leu-----GlyGluProTyrHisLysLysLysGluHisTyrLeuSer 161  
 Db 388 CTGCTGCTGCGGCCAGAGCGGCTCTGGGAC-----ACACAATATGTGGCT 432  
 QY 162 GlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAsp 181  
 Db 433 GCCAATGATCTGCGGTGACGACCGTGGGGCTG----- 465  
 QY 182 LysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPhe----- 196  
 Db 466 -----GTAGCCAGGGCTTGCGAGGGAGGCCCTCTCTTGTGGGGCGAGGATAC 516  
 QY 197 -----ProThrIleSerSerArgLysLeuThrLysAsn 207  
 Db 517 ACCAGCAGGGGTGTGGGGGTGGCATTCACCCCATCACACCGGGCCCTGTGGCGC--- 573  
 QY 208 SerGluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheValAlaSerMetIle 227  
 Db 574 CCCGACCCCAAGCTGCTTCTCTCTAT-----GAGGAGACAGCCAAAGCTG--- 618  
 QY 228 LysIleProSerAspThrPheThrIleIleProAspPheAspIleTyrValTyrGly 247  
 Db 619 -----GCAGTGGCGCGCTCTCCGAGTACAGCCACCACTCTGTGAGTGCC 663  
 QY 248 PheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMetValSerProPro 267  
 Db 664 TTTTCAGCTGGGGCCAGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 711  
 QY 268 GlySerThrThrLysGluGlnValTyrThrSerLysLeuValArgLeuCysLysGluAsp 287  
 Db 712 -----CAGGCTAGTCTAGAGCTTCTGTCCTATGTATCTCGAGTGTCTCGCGGAC 765  
 QY 238 ThrAlaPheAsnSerTyrValGluValProIleGlyCysGluArgSerGlyValGluTyr 307  
 Db 766 CAGCACTACTACTCTATGTGGAGTGTGCTCTGGCTCGCAA-----GGTGGCGGCTAC 819  
 QY 308 ArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGly 327

Db 820 GGCTGATCCAGCTCGAGCTGGCCACCTCCAGGAGGTG----- 861  
 QY 328 ValHisProAspAspLeuLeuPheThrValPheSer----- 340  
 Db 862 -----GGCATGGGAGGTCCTTTGTCAGCTTCTCCCTGGCTGCACCCCCACCTGTG 915  
 QY 341 ---LysGlyGlnLysArgLysMetLysSerLeuAspLysSerAlaLeuCysIlePheIle 359  
 Db 916 GGCAGCCGCCATCGCGGCTGTGGGCACTGCGAGCCCTGCTGCTGTGCTTCCCTCC 975  
 QY 360 LeuLysGlnIleAsnAspArgIleLysGluArgLeuGln---SerCysTyr--- 375  
 Db 976 CTGATGAGGTG---GACCGGCTGTGCTAATCGCACGCGAGATGCTCTACACCCGGAG 1032  
 QY 376 ---ArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCysSer 394  
 Db 1033 GGTCTGCTGAGGATGGACCGAGGTGGCTTACATCAGTAT---GATGCAATCTCTGAC 1089  
 QY 395 SerAlaLeuLeu-----ThrIleAspAspAsnPheCysGlyLeuAspMetAsn--- 410  
 Db 1090 TGTGCACAGCTGCAGTGGACACCTGATGCTTATCCCTGTGCTCAGACCCACACGCC 1149  
 QY 411 AlaProLeuGlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAsp 430  
 Db 1150 AGCCCATGCGCAGCGGTCCCGCTGGAAGCCACACCAATCTGGAGTGGCCAGGATT 1209  
 QY 431 ArgMetThrSerValIleAlaIleValTyrLysAsnHisSerLeuAlaPheValGlyThr 450  
 Db 1210 CAGCTAACAGCTGGGAGTCAACCATGAGATGGACACCATCCTGCTTCTGGGTGAT 1269  
 QY 451 LysSerGlyLysLeuLysLysIleArgValAspGlyProArgGlyAsnAlaLeuGlnTyr 470  
 Db 1270 AGTCAAGGCGAGTGCACAGGTCTACTTG---GGCCAGGAGCGATGGCCACCATAC 1326  
 QY 471 GluThrValGlnValAspProGlyProValLeuArgAspMetAlaPheSerLysAsp 490  
 Db 1327 TCCACAGAGCATCCACAGGCTGCTGCATGAGAGACCTCACCTTTGATGGGACC 1386  
 QY 491 HisGluGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCys 510  
 Db 1387 TTTGAGCACCTGTATGTCATGCCAGACACACTTCTGAAGTTCTCTGTGGCTTCTGT 1446  
 QY 511 GlyGlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTrpCys 530  
 Db 1447 GCTCAGCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1506  
 QY 531 ValLeuHisAsnThrCysThrArgLysGluArgCysGluArgSerLysGluProArgArg 550  
 Db 1507 GTGCTCTTGGCAGGTGTCAGTGGCGGTCTGAGTGTCTGAGGCGCCAGGCGCCAGCAG 1566  
 QY 551 -----PheAlaSerGluMetLysGlnCysValArgLeu---ThrValHisPro 565  
 Db 1567 TGGCTATGGAGCTTCCAGCCTGAGTGT---GGCTGTCTGCAAGTGGCAGCCATGATCCT 1623  
 QY 566 AsnAsnIleSerValSerGlnTyrAsnValLeuLeuValLeuGluThrTyrAsnValPro 585  
 Db 1624 GCCAACATCAGCCGAGGAGACGAGGAGGTTCCTTA-----TCAGTGCCA 1671  
 QY 586 GluLeu-----SerAlaGlyValAsnCysThrPheGluAspLeuSerGlu 600  
 Db 1672 GACCTGCCACCCCTGTGGCCAGGAGTATATCTCTGCCACTTTTGGGCAA---CATCAG 1728  
 QY 601 MetAspGlyLeuValValGlyAsnGlnIleGlnCysTyrSerProAlaAlaLysGluVal 620  
 Db 1729 AGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1788  
 QY 621 ProArgIleIleThrGluAsnGlyAspHisHisValValGlnLeuLysSerLys 640  
 Db 1789 CCA---GTCTGCCAGAGGAGCCGACTACGTATCCGTGAGCGTGGAGCTCAGATT--- 1842  
 QY 641 GluThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerVal----- 658  
 Db 1843 ---GGCGCTGTGTGATGCCCAAACTTCCCTCTCTTCTATGACTGTGTGGCGGTCACT 1899

QY 659 -----HisAsnSerCysLeuSerCysValGluSerProTyrArgCysHisTrp 674  
 Db 1900 GAATCCGCCCATCTGCGCAGTGCAGGCTGTGTGAGCAGCGCTGGGGGTGTAACCTGG 1959  
 QY 675 CysLysTyrArgHisValCysThrHis-----AspProLysThrCys 688  
 Db 1960 TGTGTCTGGCAGCACCTGTGACCCCAAGGCTGTGTGTGCTGTGGGCCCATGGTTGCA 2019  
 QY 689 SerPheGlnGlu----- 692  
 Db 2020 AGCCATCAGAGCCCTTGTTCGCCAGACCCCTCTCTGCAAGAGTGGACCCACCTCC 2079  
 QY 692 ----- 692  
 Db 2080 CCACCCACAGCCCCCAAGCCCTGGCCACCCCTGCTCTCTGACACCCCTTCCCTGGAGCCT 2139  
 QY 692 ----- 692  
 Db 2140 GGGGCTCCCTCCACAGCCACACAGCTTGGGACATCTACCTGGGGCTAGTCTTCCCTGCTC 2199  
 QY 692 ----- 692  
 Db 2200 AGCCCTGGGGCCATGGCCAGGTTCTGGCTCCATATCTTGCCCTGGCTCCACAGGTCG 2259  
 QY 692 ----- 692  
 Db 2260 CCTCTCATGAGGAGCCCTCCCTCCAGCCCCCAAAATGGACCTGGAAACCGCTGTCCCT 2319  
 QY 692 ----- 692  
 Db 2320 GCCCCACTGACTTCAGACCCCTCAGCCACACCTGAGGACCTCTTGGCCTCCCGCTGCTCA 2379  
 QY 692 ----- 692  
 Db 2380 CCGTCAGAGGTAGCAGCAGTGCCTCCCTGTCAGACCCCTGGCCCGAGGCTCTTCATCCACA 2439  
 QY 693 -----Gly 693  
 Db 2440 GTGCCCTGGACCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2499  
 QY 694 ArgValLys-----LeuProGlu----- 699  
 Db 2500 TCCGTGAAGCCCGCTGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2559  
 QY 699 ----- 699  
 Db 2560 GAGTGGACGGGGGTGACGACCCGCTTCTCCACTTCCACCTTCTCAGGTGATGGA 2619  
 QY 699 ----- 699  
 Db 2620 GACTCAGCAGAGCTTGGAGCCCTCCCGCCCTCATCTCCCTGCTCCAGCTCGACTAC 2679  
 QY 700 -----AspCysProGlnLeuLeuArgValAspLys----- 709  
 Db 2680 CAGTATGACACCCCGGCTCTGGAGCTGGAAGAGCGGACCTTGGGGGCAAGCTCCCTGC 2739  
 QY 710 -----IleLeuValProValGluValIleLysProIle 720  
 Db 2740 CCTGTGTGAGAGCGTTCAGGGCTCCAGCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCT 2799  
 QY 721 ThrLeuLysAlaLysAsnLeuProGlnSerGlyGlnArgGlyTyrGluCysIle 740  
 Db 2800 CGGCTGCTAGCAGGACCTGCACCTTTCCAGGATGGCCAGGAGACAATGAGTGTGTG 2859  
 QY 741 LeuAsnIleGlnGlySerGluGlnArgValProAlaLeu-----Arg 754  
 Db 2860 ATGAGCTGAGGCGCTCGAGGTGGTGTGAGCCCGGGTGCAGTGTGAGCCACCTCCA 2919  
 QY 755 PheAsnSerSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluTle 774  
 Db 2920 GATACCCAGTGCATACCTGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2976



5035	ACAGACACTGTGGTGAGGAAGCTGTCTACCAACTGGATGTCCATCTGTGTATACCTTC	5094
1450	LeuLysGluCysAlaGlyGluProLeuPheSerLeuPheCysAlaIleLysGlnMet	1469
5095	GTGAGGACTCCGTAGGGAGCCTCTGTACATGCTCTTTTCGAGGGATTAAGCACCAAGTG	5154
1470	GluLysGlyProIleAspAlaIleThrGlyGluAlaArgTyrSerLeuSerGluAspLys	1489
5155	GATAAGGGGCCAGTGACAGCTGTGACAGCAGAGCCAAATACACCTTGAACCAACACCG	5214
1490	LeuIleArgGlnGlnIleAspTyrLysThrLeuValLeuSer	1506
5215	CTGCTCAGAGAGATGTGAGTACCGTCCCTTGACCTTGAATGCATATTTGGCTGGGG	5274
1507	ProAspAsnAlaAsnSerProGluValProValLysIleLeuAsnCysAspThrIleThr	1526
5275	CTGGGGCAGAGAGAGCCCGCGTGTGCGCTGAAGGTCTAGACTGTGACACCATCTCC	5334
1527	GlnValLysGluLysIleLeuAspAlaIlePheLysAsnValProCysSerHisArgPro	1546
5335	CAGCAAGAGAGAGATGCTGGACCAAGCTTTAAAGGAGTGGCTCTCACCCAGCGCCA	5394
1547	LysAlaAlaAspMetAspLeuGluIleThrArgGlnGlySerGlyAlaArgMetIleLeuGln	1566
5395	GACCCTCGCACCCCTTGATGTTGAGTGGCGTCTGGGTGGCGCGGCACCTCATCTTCT	5454
1567	AspGluAspIleThrThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHis	1586
5455	GACGAGATGTCACTTCTGAGGTCAGGGTCTGTGGAGCGCCTGAACACACACGACAGC	5514
1587	TyrGlnValProAspGlySerValAlaLeuVal	1603
5515	TACAAGTCCCATGGACCACTGTGGCCCTCTCCCTGGCTCACCAAGCATGTGCTC	5574
1604	AlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSerAlaSerLysTyrGluAsn	1623
5575	CGG-----GAAAC	5583
1624	MetIleArgTyrThrGlySerProAspSerLeuArgSerArgThrProMetIleThrPro	1643
5584	CAGATTATGTCCTGGA-----GAGCGGACCCCAATGCTGGAGGAT	5625
1644	AspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHisGluHisGlyAspGln	1663
5626	GTAGATGAGGGGGCATCCGCCCTCGCACCTGGTGAAGCCAAGTAGTAGCCGCGGACCG	5685
1664	LysGluGlyAspArgGlySer-----LysMetValSer	1674
5686	CCAGCGCTCGGAGGGCAGCCTTCGGGGCGGGAGCGTGAGCGGCCAAGGCCATCCCT	5745
1675	GluIleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspAsp	1694
5746	GAGATCTACCTGACCCGCTGTCTCCATGAAGGCGACCCCTGCAGAAAGTTCGTGGATGAC	5805
1695	LeuPheGluThrIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLys	1714
5806	CTGTTCAGGTGATTTTCACACACAGCGC-----CCCGTGGCGCTCGCTGTGAAG	5856
1715	TyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisVal	1734
5857	TACTTCTTCACCTGCTGGATGAGCAGGCCCGCAGCAGCATGGCATCCGACAGGACACC	5916
1735	ArgHisThrTrpLysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsn	1754
5917	ATCCACATCTGGAAGACCAACAGCTGCCCTCTGAGTCTCTGGATCAATATAATAAANAAC	5976
1755	ProGlnPheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValVal	1774
5977	CCGAGATTGTGTCGACGTGCAAACATCTGATAACATGATGGGTGCTCTGTGTCATT	6036
1775	AlaGlnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerPro	1794
6037	GCACAGACCTTCATGGACGCTGCACCTCGCCACCAAGCTGGCGCGGAGCTCCCG	6096

Qy	1795	SerAsnLysLeuLeuTyrAlaLysAspLleProSerTyrTyrLysAsnTrpValGluArgTyr	1814
Db	6097	ATCAACAACAACTTCTGTATCCACGGGACATATCCCGGTACAAAGCGGATGGTGGAAAGGTAC	6156
Qy	1815	TyrSerAspIleGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAla	1834
Db	6157	TATGCAGACATCAGACAGACTGTCCAGCGCCAGCACCACAGAGATGAACCTGTCTGGCT	6216
Qy	1835	GlUGlnSerArgMethIleMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePhe	1854
Db	6217	GAAGTCTCTGGAACTACTCTCCGAGACCTCGGGCGCGAGTGCCTGCATGAACCTAC	6276
Qy	1855	SerTyrValGlyLysTyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCys	1874
Db	6277	AAGTACATCAACAAGTACTATGACCAAGATCATCACTGCCCTGGAGGAGGATGCACGGCC	6336
Qy	1875	GlyLysGlnLysLeuAlaTyrLysLeuGluGlnVal	1886
Db	6337	CAGAAGATGCAGCTGGGCTATCGCTCCACAGATT	6372
RESULT 5			
XX	AAH81745		
ID	AAH81745 standard; DNA; 6252 BP.		
XX	AC		
XX	AAH81745;		
XX	DT		
XX	21-SEP-2001 (first entry)		
DE	Human differential transcription-associated cDNA SEQ ID 254.		
XX	Differential transcription; human; rat; tumour cell; cytostatic;		
KW	Ras modulator; Class II tumour suppressor gene; gene therapy; ss.		
KW	Homo sapiens.		
OS	WO200157058-A2.		
XX	PN		
XX	09-AUG-2001.		
PD	31-JAN-2001; 2001WO-EP01003.		
XX	PF		
XX	31-JAN-2000; 2000DE-1004102.		
XX	PR		
XX	(META-) METAGEN GES GENOMFORSCHUNG MBH.		
PA	Rosenthal A, Hinzmann B, Schaefer R, Zuber J, Tchernitsa O;		
PI	Grips M, Hellriegel M, Schmitz A, Sers C;		
PI	WPI; 2001-483415/52.		
DR	Nucleic acids differentially expressed between tumor and normal cells,		
XX	useful for diagnosis or therapy of tumors and for screening active		
PT	agents		
PT	Disclosure; Page 403-404; 579pp; German.		
PS	This invention describes a nucleic acid (I) with differential expression		
XX	between tumour and normal cells and which has cytostatic activity. (I)		
CC	work as modulators of Ras activity by inducing expression of tumour		
CC	suppressor genes. (I), and polypeptides encoded by them, are useful as		
CC	targets for diagnosis or therapy and in screening to determine the		
CC	effects of an active compound (potential pharmaceutical) on a cell line,		
CC	particularly for diagnosis and treatment of tumors, especially by		
CC	modulating expression of (I) (by gene therapy, antisense RNA or ribozyme		
CC	methods) or by modulating the amount and/or location of (I)-encoded		
CC	polypeptides (by administration of the polypeptide or its activator,		
CC	antibody (optionally as a conjugate) or inhibitor). The method allows		
CC	identification of many Class II tumour suppressor genes (i.e. genes that		
CC	are not primary targets for tumour-initiating mutations).		
CC	AAH81492-AAH82376 represent the human and rat derived nucleic acid		
CC	fragments described in the method of the invention.		
XX			

SQ Sequence 6252 BP; 1255 A; 2022 C; 1915 G; 1060 T; 0 other;

Alignment Scores:

Pred. No.:	1.44e-202	Length:	6352
Score:	2580.00	Matches:	654
Percent Similarity:	50.63	Conservative:	344
Best Local Similarity:	33.18	Mismatches:	713
Query Match:	25.83	Indels:	260
DB:	22	Gaps:	61

US-09-964-956-13 (1-1896) x AAH81745 (1-6252)

Qy	24	ThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPhe	43
Db	35	ACCTTGCTGGCGCTGCTGGCGCGCAGGTGCCAGCTGAGGCCCGCAAGCTGGACCTTCTTC	94
Qy	44	ArgGlyGlyProAlaGluGlyPheAsnHisIleValValAspGluArgThrGlyHisIle	63
Db	95	CGCAGCGAG-----AAAGAGCTGAACCACTGGCTGTGGATGAGCGCTCAGCGCGTGGTG	148
Qy	64	TyrLeuGlyAlaValAsnArgIleTyrLysLeuSerAspLeuLysValLeuValThr	83
Db	149	TACCTGGGGCGGTGAATGCCCTCTACCAAGCTGGATGGGAAGCTGCAGCTGGAGCAGCAG	208
Qy	84	HisGluThrGlyProAspGluAspAsnProLysCysTyrProArgIleValGlnThr	103
Db	209	GTGGCCACGGCGCGCCTTGGACAACAAGAAAGTCACGCGCCCATCGAGCCAGCCAG	268
Qy	104	CysAsnGluProLeuThrThrAsnAsnValAsnLysMetLeuLeuIleAspTyrLys	123
Db	269	TGCCATGAG---GCTGAGATGACTGACAATGTCACCAAGCTGCTGCTGCAGCCCTCC	325
Qy	124	GluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLysLeuLeuArgLeu	143
Db	326	AGGAAGCGCCTGGTGGAGTGGCGACACCTCTTCAAGGGCATCTGCGCTCGCGGCCCTG	385
Qy	144	GluAsp-----LeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeuSer	161
Db	386	AGCAACATCTCCCTCCGCGCTGTTCTACGAGGACGGCAGCGGGAGAAAGTCTTTCGGTGGCC	445
Qy	162	GlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAsp	181
Db	446	AGCAATGATGAGGGCGTGGCCACAGTAGTGGGGCTGGTGAGCTCCACGGGTCTCGTGGTGAC	505
Qy	182	LysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSer	201
Db	506	CGCGTGCTGTTGTGGGCAAGGCAATGGCCACAGCACAGCGCATCATCGTGAGCACT	565
Qy	202	ArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGlu	221
Db	566	CGGCTGTTGGACCGGACTGACAGCAGGGAGCGCTTTGAAGCCTACAGGACCACCGCCACC	625
Qy	222	PheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAsp	241
Db	626	TACAAGCGCGCTACCTGTCACCAACACACAGCAGTTC-----	664
Qy	242	IleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnPro	261
Db	665	-----GTGGCGGCCTTCGAGGACGGCCCTACGTCTCTTGTCTTCACACAGCAG	715
Qy	262	GluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuVal	281
Db	716	GACAAG---CACCGCGCCGGAACCGCAGC-----	751
Qy	282	ArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCysGlu	301
Db	752	CGCATGTGCAGAGAAGACCCCAACTACTCTACTCTGGAGATGCACCTGCAGTGCCTGG	811
Qy	302	ArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaVal	321
Db	812	GACCCCGACATCCAC-----GCCCTGCTCTT-----	847
Qy	322	LeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerLys	341

1895 TCCTGGTGGACCAACCGCTGGACCTGCCAGTGGGACCTGCCTACCAAGAGTGGCGGGAG 1954  
QY ---AspProLysThrCysSerPheGlnGluGly-----ArgValLysLeuProGluAsp 700  
1955 GCTTCCGCCCAACCGCT-----GAGGACGGCATCGCTGCCCATATGGAGACAGC 2005  
QY CysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysProIle 720  
2006 TGTCGCCAGTTCCTGGGACCGCCCGCTGGTATCCCATGAACACGACGACAGATGTG 2065  
QY ThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrgluCysIle 740  
2066 AACTTCAGGGCAAGAACCTGGAC----- 2089  
QY LeuAsnIleGlnGlySerGluGlnArgValProAla-----LeuArgPheAsnSerSer 758  
2090 ---ACCGTGAAGGTTCTCCCTGCACGTGGGACGTGACTTGCTCAAGTTTCATGGAGCCG 2146  
QY SerValGlnCysGlnAsnThrSerTyrrSerTyrrGluGlyMetGluIleAsnAsn----- 776  
2147 GTGACCATGCAGGAATCTGGGACCTTCGCCCTTCGGACCCCAAGACTGTCGCCAGATGCC 2206  
QY -----LeuProValGluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsn 793  
2207 AACGAGACGCTGCCCTGCACCTACGTCGAAGCTTACGGCAAG---AATATCGAC--- 2260  
QY ProAlaGlnAsnLysValHis-----LeuTyrrLysCysGlyAlaMetArgGluSerCys 811  
2261 -----AGCAAGCTCCATGTGACCTCTACAACTGCTCTTTGGCCGCGGACGACTGC 2311  
QY GlyLeuCysLeuLysAlaAspPheAlaCysGlyTyrrCysGlnGlnProGlyGln 831  
2312 AGCTGTGGCGGCGGCTACCCCGACTACAGGTGGTGGTGGGGGCCGAGGACG 2371  
QY CysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuLeuSerGlyAla 851  
2372 TGCCTGTATAGGCCCTGTGC-----AACACC 2398  
QY LysSerLysCysThrAsnProArgIleThrGluIleIleProValThrGlyProArgGlu 871  
2399 ACCTCCAGTGGCGCGCGCGCTCATCCAGGATCCAGCTGACGCGGCGCCCTGGGT 2458  
QY GlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPheArgAspIleAla 891  
2459 GGGGCAATCCGATCACCATCCTGGGTCCCAATTGGGCGTCCAAGACGAGGACATC--- 2515  
QY SerHisValLysValAlaGlyValGlyCysSerProLeuValAspGlyTyrrIleProAla 911  
2516 CAGAGGATCTCTGGCGCGCGGAACTGCTCTTCAGCCGGAACGTTACTCGGTGCC 2575  
QY GluGlnIleValCysGluMetGlyGluAlaLysProSerGlnHisAlaGlyPheValGlu 931  
2576 ACCCGATCGTGTGTGATC---GAGGCTCGGAGACGCGCTTTCACGGGGGTGTCGAG 2632  
QY IleCysValAlaValCysArgProGluPhe-----MetAlaArgSerSer----- 946  
2633 GTGACGTC-----TTCGGGAACTGGCGCGCTTCGCCCTCCCAATGTC 2674  
QY GlnLeuTyrrPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMet 966  
2675 CAGTTCACCTTCCACAGCCCAAGCCTCTCAGT---GTGGAGCCCGACGAGGACCGGAG 2731  
QY SerGlyGlyThrGlnValThrIleThrAsnLeuAsnAlaClySerAsnValVal 986  
2732 CGCGGGCCACACATCACCATCCACGCGCACCCACCTGGACGCGGCTCCCGAGGAGC 2791  
QY ValMetPheGlyLysGlnProCysLeuPheHisArgArgSerProSerTyrrIleValCys 1006  
2792 GTG-----CGGGTGACCTCAACGGCGTCCCGTGT 2821  
QY AsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnValAspArgAla 1026  
2822 AAGTGACCAAGTTGGGGCGCAGCTCAGTGTGTCTACTGGCCCCCGGACGCGGCGC 2881

1027 LysIle-----HisGlnAspLeuValPhe 1034  
QY -----  
Db CAGATGCTTCTGGAGGCTCTCTACGGGGGTCCTCCCGTCCCAACCGCGCATCTTCTTC 2941  
QY GlnTyrrValGluAspProThrIleValArgIleGluProGluTrpSerIleValSerGly 1054  
2942 ACCTACCGCGAAACCCCGTACTCGAGCGCTTTCGAGCGCTACGAAGCTTTGCGACGTGT 3001  
QY AsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArg 1074  
3002 GCGCGCACATCAACGTACGGGTTCAGGCTTCAGCTCATCAGAGGTTTGCATG--- 3058  
QY AlaLysHisGlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMet 1094  
3059 -----GTGGTCACTCGCGGACCCCTGCAGTCTCTGGCAGCGC 3094  
QY ThrCysGlnAla-----ProAlaLeuAlaLeuLeuProAspHis----- 1107  
3095 CCGCGGAGGCTGAATCCCTGCAGGCCATGAGGGTGGTGGTACAGACTACGTTGTCCAC 3154  
QY PheIle-----GlnSerAspLeuThrGluArgProGluGluPheGly 1119  
3155 AATGACACCAAGTGTCTTCTGTCCTCCCGGCTGTCCTGAGGAGCCAGAGGCTTACAAC 3214  
1120 PheIle-----LeuAsnValGlnSerLeuLeuIleLeuAsnLysThrAsn 1135  
3215 CTCAGGTGCTGATCGAGATGCGCGGACCGCTGCTCTCAGAACAGAGGCGCGGCC 3274  
1136 PheThrTyrrTyrrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeu 1155  
3275 TTCGAGTACGTGCTGACCCACCTTTGAGAACTTC----- 3310  
QY LysProGlyThrProIleIleLeuLysGlyLysAsnLeuIleProValAlaGlyGly 1175  
3311 -----ACAGTGGCGTCAAGAGCAGGTCACAGCTCATCCACGCGCGGGGACC 3361  
1176 AsnValLysLeuAsnTyrrThrVal-----LeuValGlyLysProCys 1190  
3362 AATCTGAACAGCGCATGACGTGACGAGGCGGAGGCGCTTCTGGTGGCGGAGCGTGC 3421  
QY ThrVal---ThrValSerAspValGlnLeuLeuCysGluSerProAsnLeuIleGlyArg 1209  
3422 ACCATGAAGACGCTGACGGAGACCGACCTGTACTGTGAGCCCCCGAGGTGACGCGCCG 3481  
1210 HisLys-----ValMetAlaArgVal 1216  
3482 CCCAAGCGCGCAGAAACGAGACACACACACCTGCCGCGATTCATTGTGAAGTTC 3541  
1217 GlyGlyMetGluTyrrSerProGlyMetValTyrrIle-----AlaProAspSerPro 1233  
3542 GGCTCTCGCGAGTGGGTGCTGGCGCGGTGAGTACACACACGCGGTGACGCGGTGCGG 3601  
QY LeuSerLeu-----ProAlaIleValSerIleAlaValAlaGly 1246  
3602 CTCAGCTCATCTGCGCGTGGTGCATCGTCCCATCGTGTGCTCATCGCGGTCTCT--- 3658  
1247 GlyLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrrLysArgLysSerArgGlu 1266  
3659 -----GTCTACTGCTACTTGGAGGAAGAGCCAGCAG 3688  
1267 SerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAla 1286  
3689 GCGACAGAGATGATGAGAAGATCAAGTCCACCTGGAGGCGCTGGAGGAGAGCGGTGCGG 3748  
1287 LeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIleHisGlnLeuThrSerAsp 1306  
3749 GACCGCTGCAAGGAAGAAATCACAGCTCATCATCGAGTGGAGGACGAGCAGCAGC 3808  
1307 LeuAspGlyAlaGlyIleProPheLeuAspTyrrArgThrTyrrMetArgValLeuPhe 1326  
3809 GTGCACGAGCGCGGCATCCCGTGTGGTGGTACTAGAAGACCTACACCGACCGCGCTTCTTC 3868

QY 1327 -----ProGlyIleGluAspHisProValLeuArgAspLeuGluValProGly 1342  
 DB 3869 CTGCCCTCCAGGAGCGGCAAGAGCGTGTATGATCACCAGCAAGCTGCACATCCCTGAG 3928  
 QY 1343 TyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIleAsnLys 1362  
 DB 3929 CGCGCGCGCGGTGGTGGAGCAGCCCTCTACCAAGTTCCTCAACCTGTGTGAACAGCAAG 3988  
 QY 1363 ValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPheSerMetArgAsp 1382  
 DB 3989 TCTTTCTCTCATCAATTTCTACACACCTCTGGAGAACACAGCGGAGTTCCTCGCCCGCGCC 4048  
 QY 1383 ArgGlyAsnValAlaSerIleMetThrValLeuGlnSerLysLeuGluTyrAlaThr 1402  
 DB 4049 AAGGTCTACTTCCGCTCCCTGCTGCGGTGGCGCTGCACGGGAAACTGGAGTACTACACG 4108  
 QY 1403 AspValLeuLysGlnLeuAlaAspLeuIleAspLysAsnLeuGluSerLysAsnHis 1422  
 DB 4109 GACATCATGACACACCTCTCTCTGAGCTCTCTGGAGCAGTACGTGGTGCCCAAGAAC--- 4165  
 QY 1423 ProLysLeuLeuArgArgThrGluSerValAlaGluLysMetLeuThrAsnTrpPhe 1442  
 DB 4166 CCCAAGCTGATCTCGCAGGTCTGAGACTGTGTGGAGAGGATGCTGTCCAACCTGGATG 4225  
 QY 1443 ThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeuPheSerLeuPhe 1462  
 DB 4226 TCCATCTCGCTGTACCACTACCTCAAGGACAGTGGCGGGAGCCCTGTGTACAAAGCTCTTC 4285  
 QY 1463 CysAlaIleLysGlnMetLysGlyProIleAspAlaIleThrGlyGluAlaArg 1482  
 DB 4286 AAGGCCATCAACATCAGTGTGAAGGCGCGGTGGATGCGGTACAGAAGAAGGCCAAG 4345  
 QY 1483 TyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLysThrLeuValLeu 1502  
 DB 4346 TACACTCTCAACGACACAGCGGTCTGTGGGGATGATGTGGAGTACGACCCCTGACGGTG 4405  
 QY 1503 SerCysValSerProAspAsnAlaAsnSerProGluValProValLysIleLeuAsnCys 1522  
 DB 4406 AGCGTGTGCTGCAGCAGCAG---GGAGTGGACGCCATCCCGGTGAAGCTCTCAACTGT 4462  
 QY 1523 AspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLysAsnValProCys 1542  
 DB 4463 GACACCATCTCCAGGTCAAGGAGAGATCATGTACCAAGTGTACCGTGGGCGAGCCCTGC 4522  
 QY 1543 SerHisArgProLysAlaAlaAspMetAspLeuGluTyrArgGlnGlySerGlyAlaArg 1562  
 DB 4523 TCCTGTGGCCCGCAGCCAGACAGCGTGTCTGGAGTGGCGTCCGGCTCCACAGCGCAG 4582  
 QY 1563 MetIleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTrpLysArgLeuAsn 1582  
 DB 4583 ---ATCCTGTGGACCTGGACCTGACGTACAGCGGGGCGGTGGAGCGCGTCAAC 4639  
 QY 1583 ThrLeuAlaHisTyrGlnValProAspGlySerValValAlaLeuValSerLysGlnVal 1602  
 DB 4640 ACCCTTATGCACTACAATGTCGGGATCGAGCCACCTCATCTG----- 4684  
 QY 1603 ThrAlaTyrAsnAlaValAsnSerThrValSerArgThrSerAlaSerLysTyrGlu 1622  
 DB 4685 -----TCCAAGTGGGGTCTCCACGACGCGGGAGGACAGCCAGCAG 4726  
 QY 1623 AsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgThrProMetIleThr 1642  
 DB 4727 GACCTG-----CCTGGGAGCGCCATGCCCTC----- 4753  
 QY 1643 ProAspLeuLysGlyValLysMetTrpHisLeuValLysAsnHisGluHisGlyAsp 1662  
 DB 4754 -----CTGGAGGAGGAGACCGGTGTGGCACTGTGTCGCGCCGACAGCGGTGGAC 4807  
 QY 1663 GlnLysGluAspArgGlySer-----LysMetValSerGlu 1675  
 DB 4808 GAGGCAAGTCCAAGAGAGCGCGTGAAGAGAGAGGAGCGGACGAGGAGCCATCACCAG 4867  
 QY 1676 IleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspLeu 1695

DB 4868 ATCTACCTGACGGGCTGCTCTAGTCAAGGGCACACTGCAGCAGTTTGTGGACAACATTC 4927  
 QY 1696 PheGluThrIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyr 1715  
 DB 4928 TTCAGAGCGTGTGGCGCCTGGGCAC-----CGGTGGCCACCTGCAGTCAAGTAC 4978  
 QY 1716 MetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArg 1735  
 DB 4979 TTCTTCGACTTCTCTGGAGCAGCAGCAGGAGAGACACAACTCCAGGATGAAGACACATC 5038  
 QY 1736 HisThrTrpLysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnPro 1755  
 DB 5039 CACATCTGGAGACGAACAGCTTACCGCTCGGTCTGTGGTGAACATCTCTCAAGAACCCT 5098  
 QY 1756 GlnPheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValValAla 1775  
 DB 5099 CACTTCATCTTTCAGCTGCATGCCAGGAGTGGTGACGCGCTCGTGTAGTATCCTCCG 5158  
 QY 1776 GlnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSer 1795  
 DB 5159 CAGACCTTCATGGATGCTGCACGCGCAGGAGCATAAAGCTGAGCGCGATCTCTCCACAG 5218  
 QY 1796 AsnLysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrTyr 1815  
 DB 5219 AACAAAGCTGTGTACGCCAAGGAGATCTCCACCTACAAAGAGATGGTGGAGGATTACTAC 5278  
 QY 1816 SerAspIleGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGlu 1835  
 DB 5279 AAGGGATCCGGCAGATGGTGCAGTCAAGCAGCAGGAGCATGACACACACTTGCACAG 5338  
 QY 1836 GlnSerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSer 1855  
 DB 5339 ATTTCCCGCGCGCACACGAGTCTCTTGAACACCTCTGTGGCAGCTCCACACGCTCTACCAA 5398  
 QY 1856 TyrValGlyLysTyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCysGly 1875  
 DB 5399 TACACGCAAGAGTACTATGACGAGATCATCAATCCCTTGGAGGAGGATCCTTGCCTGCC 5458  
 QY 1876 LysGlnLysLeuAlaTyrLysLeuGluGlnVal 1886  
 DB 5459 AAGATGACGTGGCGCTTCCGCTGCAGCAGATT 5491  
 RESULT 6  
 AAS00020  
 ID AAS00020 standard; cDNA; 6252 BP.  
 XX AC AAS00020;  
 XX DT 09-MAY-2001 (first entry)  
 XX DE Human cDNA encoding Plexin-B2.  
 DE Human; Plexin-B2; semaphorin domain; hyperplasia; neoplasia; cancer;  
 KW neurodegenerative disease; autoimmune disease; lupus; multiple sclerosis;  
 KW inflammatory bowel disease; diabetes type I; rheumatoid arthritis;  
 KW immunogen; antibody; ss.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 FT CDS 11..5527  
 FT /\*tag= a  
 FT /product= "Plexin-B2"  
 FT sig\_peptide 11..64  
 FT /\*tag= b  
 FT mat\_peptide 65..5524  
 FT /\*tag= c  
 FT /label= "Mature Plexin-B2"  
 XX WO200114420-A2.  
 XX PD 01-MAR-2001.

XX -25-AUG-2000; 2000WO-US23365.  
 XX GlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAsp 181  
 PR 25-AUG-1999; 99US-0150576.  
 XX AGCAATGATGAGGGCTGGCCACAGTGGGCTGGTGGAGCTCCACGGGTCTGCTGGTGAC 505  
 PA (UYTO-) UNIV TORINO.  
 PA (REGC ) UNIV CALIFORNIA.  
 XX LysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSer 201  
 XX CGCGTGTGTTGTGGCAAGGCAATGGGCCACACGACACGACGACGACGACGACGAC 565  
 PI Artigiani S, Comoglio PM, Goodman CS, Tesier-Lavigne M;  
 PI Tamagnone L;  
 XX WPI: 2001-226610/23.  
 DR P-PSDB; AAU00015.  
 DR New plexin polynucleotides and polypeptides, useful in diagnosis,  
 PT therapy and in producing compounds for treating diseases involving  
 PT aberrant cell growth (e.g. cancer) or immune regulation (e.g.  
 PT autoimmune diseases)  
 XX Claim 2; Page 50-52; 79pp; English.  
 XX The sequence encodes Human Plexin-B2. Plexins are large transmembrane  
 CC proteins whose extracellular domain shares homology with Scatter factor  
 CC receptors and contain an approximately 500 amino acid Semaphorin domain.  
 CC The plexin polynucleotides and polypeptides, and plexin-specific binding  
 CC agents are useful in diagnosis, therapy and in the biopharmaceutical  
 CC industry. In particular, the plexin polynucleotides and polypeptides are  
 CC useful for generating compounds (e.g. plexin-specific binding agents or  
 CC antibodies) for treating or diagnosing a disease or disorder involving  
 CC aberrant cell growth (e.g. hyperplasia, neoplasia, cancer or  
 CC neurodegenerative disease), or diseases or disorders involving aberrant  
 CC immune regulation (e.g. autoimmune diseases such as lupus, inflammatory  
 CC bowel disease or Diabetes Type I), or immunosuppressive diseases such as  
 CC multiple sclerosis or rheumatoid arthritis.  
 XX SQ Sequence 6252 BP; 1255 A; 2022 C; 1915 G; 1060 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.44e-202 Length: 6252  
 Score: 2580.00 Matches: 654  
 Percent Similarity: 50.63% Conservative: 344  
 Best Local Similarity: 33.18% Mismatches: 713  
 Query Match: 25.83% Indels: 260  
 DB: 22 Gaps: 61

us-09-964-956-13 (1-1896) x AAS00020 (1-6252)

Qy 24 ThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPhe 43  
 Db 35 ACCGTGCTGGCCCTGCTGGCGCGAGGTCCAGCCTGAGGCCGCCGCCGCAAGCTGACTTCTTC 94  
 Qy 44 ArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIle 63  
 Db 95 CGCAGCAG-----AAGAGCTGAACACCTGGCTGGTGGATGAGGCCTCAGCGCTGTGTG 148  
 Qy 64 TyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThr 83  
 Db 149 TACCTGGGGGGGTGAATGCCCTTACCAGCTGGATGCGAAGCTCAGCTGGAGCAGCAG 208  
 Qy 84 HisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIleValGlnThr 103  
 Db 209 GTGGCCAGCGCGCGCCCTGGACAAACAAAGAGTCCAGCCGCCCTCATCGAGCCAGCCAG 268  
 Qy 104 CysAsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIleAspTyrLys 123  
 Db 269 TGCCATGAG---GCTGAGATGACTGACAATGTCAACAGCTGCTGCTGCTCGACCTCC 325  
 Qy 124 GluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeuLeuArgLeu 143  
 Db 326 AGGAAGCCCGCTGGTGGAGTGGGCGAGCTCTTCAAGGGCATCTGCTCTGCGCGCCCTG 385  
 Qy 144 GluAsp-----LeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeuSer 161  
 Db 1432

Db 386 AGCAACATCTCCCTCCGCTCTTCTACGAGCAGCAGCAGCGGGAGAAAGTCTTTTCGTGGCC 445  
 Qy 162 GlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAsp 181  
 Db 446 AGCAATGATGAGGGCTGGCCACAGTGGGCTGGTGGAGCTCCACGGGTCTGCTGGTGAC 505  
 Qy 182 LysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSer 201  
 Db 506 CGCGTGTGTTGTGGCAAGGCAATGGGCCACACGACACGACGACGACGACGACGAC 565  
 Qy 202 ArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGlu 221  
 Db 566 CGGTGTGGACCGGACTCAGCAGGAGGAGGCTTTGAAGCCTTACAGCAGCAGCAGCAGC 625  
 Qy 222 PheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAsp 241  
 Db 626 TACAAGCGCGCTACTCTGCCACCAACACACAGCAGCTTC-----664  
 Qy 242 IleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnPro 261  
 Db 665 -----GTGGCGCTTCGAGGAGCGGCCCTACTCTCTTCTTCTTCAACACGACG 715  
 Qy 262 GluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuVal 281  
 Db 716 GACAAG---CACCCGCGCCGGAACCCGACG-----CTGCTGGCA 751  
 Qy 282 ArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCysGlu 301  
 Db 752 CGCATGTGCAGAGAAGACCCCACTACTCTCTACCTGGAGTGGACCTGAGTGGCGG 811  
 Qy 302 ArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaVal 321  
 Db 812 GACCCCGCATCCAC-----GCCGTGCGCTT-----GGCCTGC 847  
 Qy 322 LeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerLys 341  
 Db 848 CTGGCGCGCTCGCTGGCTGGCTGGCAGGTGCTATATGCTGCTTTCAGCAGA 907  
 Qy 342 GlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLys 361  
 Db 908 GACAGC-----CGGAGCAGTGGGGCGCGCTGGCGGCTCTGCTGTGTTCCCGCTGG 961  
 Qy 362 GlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThrLeu 381  
 Db 962 AAGGTGCACGCCAAGATGGAGGCCAACCGCAGCCTGTTC-----ACAGGCACCCG 1015  
 Qy 382 AspLeuAlaTrpLeuLysValLys-----AspIleProCysSerSerAlaLeu 397  
 Db 1016 GAGCGCGCTGACATCTTCTACAAGCCCTTCCAGCGGATATCCAGTGGCGCGCCAGCG 1075  
 Qy 398 LeuThrIleAspAsnPhe---CysGlyLeuAsp---MetAsnAlaProLeuGlyVal 415  
 Db 1076 CGGGCTCCAGCAAGAGCTTCCCATGTGGCTCGAGCACCCTGCCCTACCCGCTGGGAGC 1135  
 Qy 416 SerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThrSerVal 435  
 Db 1136 CGCGAGCGCTCAGAGCAGCAGCGCTGCTCCAGCTGGAGCGCTGAACCTCAGCGCGCTG 1195  
 Qy 436 IleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeu 455  
 Db 1196 ACGGTGCGCGCGCAGAACACACCTGTTGCTTTCTGGGCACTCTGATGGCGCGATC 1255  
 Qy 456 LysLysIleArgValAspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGlnVal 475  
 Db 1256 CTCAGGTGTACCTACCTACC---CCAGTGGCACCTCTCTCAGAGTACGACTCTATCTTGTG 1312  
 Qy 476 ValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuTyr 495  
 Db 1313 GAGATTAACAAGAGAGTCAAGCGCGACCTGTGTCTGTGGAGAGCTGGGAGCGCTGTAC 1372  
 Qy 496 IleMetSerGluArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrGlnSer 515  
 Db 1373 GCCATGACCCAGCAGCAAGGTGTTCCGCTGCGAGGAGTGCCTGAGCTATCCCGACC 1432

516 CysGlyGluCysLeuGlySerGlyAspProHisCysGlyTrpCysValLeuHisAsnThr 535  
 1433 TGCACCAAGTCCGCGACTCCACGAGACCCCTACTCGCGGTGTGCTCGAGGGACGA 1492  
 536 CysThrArgLysGluArgCysGluArgSerLysGluProArgArgPheAla---SerGlu 554  
 1493 TGCACCGGAGGCGAGTGTCCGCGGCGGAGGAGCCAGCCACTGGTGTGGACCGGA 1552  
 555 MetLysGlnCysValArgLeuThr---ValHisProAsnAsnIleSer---ValSerGln 572  
 1553 AGCAAGTCTGTGGCGCTACCAAGCGCCAGCCACAGAAATGAGCCGCGGGGCCAG 1612  
 573 TyrAsnValLeuValLeuGluThrTyrAsnValProGluSerAla-----Gly 590  
 1613 GGGGAGGTGCAGCTGACCGTCAAGCC-----CTCCCTGCCTGAGCGAGGAGGACGAG 1666  
 591 ValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValGlyAsnGlnIle 610  
 1667 TTGCTGTGCTTTTGGGAGTGGCGGCCACACCCGCGCGGTGGAGGGCGAGCCGTC 1726  
 611 GlnCysTyrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGly---Asp 629  
 1727 ATCTCAACTCCCAAGCAGC-----ATCCCC-----GTCAACCGCCAGCCAGGAC 1774  
 630 HisHisValValGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerThr 649  
 1775 CACGTGGCGGTGACCATCCAGCTCCTCTTAGACGAGCAACATCTCTCAGCTCTAC 1834  
 650 SerPheValPheTyrAsnCysSerValHisAsnSer-----CysLeu 663  
 1835 CAGTACCCCTTACGACTGCGCCAGGCGCATGAGCCTGGAGAGAACCTGCGTGCATC 1894  
 664 SerCysValGluSerProTyrArgCysHisTyrCysLysTyrArgHisValCysThrHis 683  
 1895 TCCTGGGTGACCAACCGTGGACCTGCCAGTGGACCTCGGTACCAACAGTCCGGGAG 1954  
 684 ---AspProGlyThrCysSerPheGlnGluGly-----ArgValLysLeuProGluAsp 700  
 1955 GCTTGGCCCAACCCCT-----GAGGACGGCATGCTCGTCCGCACATGGAGGACAGC 2005  
 701 CysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysProIle 720  
 2006 TGTCCCAAGTCTCTGGGACCCAGCCCGCTGGTATCCCTCCCAACCAACGAGACAGATGG 2065  
 721 ThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCysIle 740  
 2066 AACTTCCAGGGCAAGACCTGCAC-----2089  
 741 LeuAsnIleGlnGlySerGluGlnArgValProAla-----LeuArgPheAsnSerSer 758  
 2090 ---ACCGTGAAGGGTTCTCTCCCTGACGTGGGCGAGTGTGCTCAAGTTTCATGGAGCCG 2146  
 759 SerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetLeuLeuAsnAsn----- 776  
 2147 GTGACCATGCAAGAACTCTGGGACCTTCGCTTCCGAGACCCCAAGAGCTGTCCACAGATGC 2206  
 777 -----LeuProValGluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsn 793  
 2207 AACGAGACGCTCCCTCTGACCTCTACGTCAAGTCTTACGGCAAG---AATATCGAC--- 2260  
 794 ProAlaGlnAsnLysValHis-----LeuTyrLysCysGlyAlaMetArgGluSerCys 811  
 2261 -----AGCAAGCTCATGTGACCCCTACAACTGTCTCTTTGGCCGCGCAGGACTGC 2311  
 812 GlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTrpCysGlnGlyProGlyGln 831  
 2312 AGCCTGTGCGGCGCGCTCAACCCGACTACAGGTGTGCGTGTGGGGGCCAGACAGG 2371  
 832 CysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuGluLeuSerGlyAla 851  
 2372 TCCGTGTATGAGCCCTGTGC-----AACACC 2398

852 LysSerLysCysThrAsnProArgIleThrGluIleLeuProValThrGlyProArgGlu 871  
 2399 ACCTCCGAGTCCCGCGCGCGTCCATCACCAGGATCCAGCCTGAGACGGGCCCTGGGT 2458  
 872 GlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPheArgAspIleAla 891  
 2459 GGGGCGATCCGCATCACCATCTCTGGGTCCAAATTTGGCGCTCCAAAGCAGGGGACATC--- 2515  
 892 SerHisValLysValAlaGlyValGluCysSerProLeuValAspGlyTyrIleProAla 911  
 2516 CAGAGGATCTCTGTGGCGCGCGGAACCTCTCTTTCAGCCGGAACGTTACTCCTGTGCC 2575  
 912 GluGlnIleValCysGluMetGlyGluAlaLysProSerGlnHisAlaGlyPheValGlu 931  
 2376 ACCGGATCGTGTGTGATC---GAGGCTCGGAGACGCTTTCACGGGGGTGTGCGAG 2632  
 932 IleCysValAlaValCysArgProGluPhe-----MetAlaArgSerSer----- 946  
 2633 GTGACACGTC-----TTCGGGAACCTGGGCGGTTCGCTTCCCTCCCAATGTC 2674  
 947 GlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMet 966  
 2675 CAGTTCACTCCCAACAGCCCAAGCCTCTCACT---GTGAGCCGACACGAGGACCGCAG 2731  
 967 SerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnValVal 986  
 2732 CGGGCGGCGCACCATGACCATCCAGGCAACCCACCTGGACACGGGCTCCAGGAGGAC 2791  
 987 ValMetPheGlyLysGlnProCysLeuPheHisArgArgSerProSerTyrIleValCys 1006  
 2792 GTG-----CGGTGACCCCTCAACGGCGTCCCGTGT 2821  
 1007 AsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnValAspArgAla 1026  
 2822 AAGTGACGAAGTTGGGCGCCACCTCCAGTGTCTACTGCCCCCGCCAGGACAGGGGC 2881  
 1027 LysIle-----HisGlnAspLeuValPhe 1034  
 2882 CAGATGTTCTGGAGGTCTCTAGGGGGGTCCCGCTGCCCAACCCCGCATCTCTTC 2941  
 1035 GlnTyrValGluAspProThrIleValArgIleGluProGluTrpSerIleValSerGly 1054  
 2942 ACCTACCGCAAAACCCGCTACTGCGAGCCTTCAGAGCCGCTACGAAGCTTTGCCAGTGT 3001  
 1055 AsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArg 1074  
 3002 GGCGCGCAGCATCAACGTACGCGGTTCAGGCTTCAGCTGATCCAGAGGTTTCCCATG--- 3058  
 1075 AlaLysHisGlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMet 1094  
 3059 -----GTGTCATCGGAGCCCTGCAGTCTCCAGTCTCCGAGCGC 3094  
 1095 ThrCysGlnAla-----ProAlaLeuAlaLeuGlyProAspHis----- 1107  
 3095 CCGCGGAGGCTGAATCCCTGCAGCCCATGACGGTGGTGGGTACAGACTAGCTGTTCAC 3154  
 1108 -----GlnSerAspLeuThrGluArgProGluPheGly 1119  
 3155 AATCACACCAAGTGTCTTCTCTCCCGGCTGTGCTGAGGACCCAGAGCCCTACAC 3214  
 1120 PheIle-----LeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsn 1135  
 3215 CTCACGCTGTGATCGAGATGAGCGGCGCACCGTGCCTGCTCAGAACAGAGGCGCGGCC 3274  
 1136 PheThrTyrTyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeu 1155  
 3275 TTCAGTACGTGCTGCTGACCCACCTTTGAGAATTC-----3310  
 1156 LysProGlyThrProIleIleLeuLysGlyLysAsnLeuIleProProValAlaGlyGly 1175  
 3311 -----ACAGTGGCGTCAAGACGAGTCAACAGCTCATCCAGCCCGCGGGGACCC 3361  
 1176 AsnValLysLeuAsnTyrThrVal-----LeuValGlyGluLysProCys 1190

Db 3362 AATCTGAACAGCGGATGACGCTGCAGGAGCGGAGCCCTTCGTGGTGCAGCGCTGC 3421  
 QY 1191 ThrVal---ThrValSerAspValGlnLeuLeuCysGluSerProAsnLeuIleGlyArg 1209  
 Db 3422 ACCATGAGACGCTGACGAGACCGACCTGCTACTGTGAGCCCGCGGAGCTGACGCCCGC 3481  
 QY 1210 HisLys-----ValMetAlaArgVal 1216  
 Db 3482 CCCAAGCGCGGAGAACAGACACACACACACCTGCCGAGTTCATTTGTAAGTTC 3541  
 QY 1217 GlyGlyMetGluTyrSerProGlyMetValTyrIle-----AlaProAspSerPro 1233  
 Db 3542 GGCTCCGCGAGTGGTGGCGCGGTGAGGTACGACACACACGCGGTGAGCGAGCGTGC 3601  
 QY 1234 LeuSerLeu-----ProAlaIleValSerIleAlaValAlaGly 1246  
 Db 3602 CTCAGCTCATCTTCCGCTGGTGCATCGTGCCATGTTGGTGCATCGCGGTCT--- 3658  
 QY 1247 GlyLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrLysArgLysSerArgGlu 1266  
 Db 3659 -----GTCTACTGCTACTGAGGAGGAGGAGCGAGCAG 3688  
 QY 1267 SerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAla 1286  
 Db 3689 GCCAAGCAGAGATGAGAAGATCAAGTCCAGCTGGAGGCGCTGGAGGAGCGTGC 3748  
 QY 1287 LeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIleHisGluLeuThrSerAsp 1306  
 Db 3749 GACCGCTGCAAGAGGAATTCACAGACCTGATGATCGAGATGGAGGAGGAGCACCACGAC 3808  
 QY 1307 LeuAspGlyAlaGlyIleProPheLeuAspTyrArgThrTyrMetArgValLeuPhe 1326  
 Db 3809 GTGCACGAGCGCGCTCCCGTGTGCTGACTACAGACCTACACCGCGCTCTCTTC 3868  
 QY 1327 -----ProGlyIleGluAspHisProValLeuArgAspLeuGluValProGly 1342  
 Db 3869 CTGCCCTCCAAAGGAGCGGACAGGAGCTGATGATCACCGGCAAGCTGGACATCCCTGAG 3928  
 QY 1343 TyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIleAsnLys 1362  
 Db 3929 CCGCGCGCGCGGTGGAGCAGCGCCCTCTACCACTGCTCCAACCTGCTGAACAGCAAG 3988  
 QY 1363 ValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPheSerMetArgAsp 1382  
 Db 3989 TCTTCTCTCATCAATTTCCATCCACACCTGGAACACCGGAGTTCCTCGCCCGCGCC 4048  
 QY 1383 ArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeuGluTyrAlaThr 1402  
 Db 4049 AAGGTCTACTTCCGCTCCCTGCTGAGGTGGCGCTGCACCGGAACTGGAGTACTACAG 4108  
 QY 1403 AspValLeuLysGlnLeuAlaAspLeuIleAspLysAsnLeuGluSerLysAsnHis 1422  
 Db 4109 GACATCATGCACCGCTCTCTCTGAGGCTCTCTGGAGCAGTACGTGTGGCCAAAGAAC--- 4165  
 QY 1423 ProLysLeuLeuLeuArgThrGluSerValAlaGluLysMetLeuThrAsnTrpPhe 1442  
 Db 4166 CCCAAGCTGATGCTGCGCAGGTCTGAGACTGTGGTGAGGAGTGTGTCACAGCTCTTC 4225  
 QY 1443 ThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeuPheSerLeuPhe 1462  
 Db 4226 TCATCTGCTGTACAGTACCTCAAGGACGTGCGCGGGAGCCCTGTACAGCTCTTC 4285  
 QY 1463 CysAlaIleLysGlnMetGluLysGlyProIleAspAlaIleThrGlyGluAlaArg 1482  
 Db 4286 AAGGCCATCAACATCAGTGGGAAAGGCGCGGTGATGCGGTACAGAGAGAGGCAAG 4345  
 QY 1483 TyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLysThrLeuValLeu 1502  
 Db 4346 TACACTCTCAACGACACGCGGTGCTGGGGATGATGTGGAGTACGACCCCTGACGGTG 4405  
 QY 1503 SerCysValSerProAspAsnAlaAsnSerProGluValProValLysIleLeuAsnCys 1522

Db 4406 AGCGTATGCTGTCAGGACGAG---GGAGTGGAGCCATCCCGGTGAAGTCTCTCAACTGT 4462  
 QY 1523 AspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLysAsnValProCys 1542  
 Db 4463 GACACCATCTCCAGTCAAGGAGAGATCATTTGACAGGTGTACCGTGGCGGACCGCTGC 4522  
 QY 1543 SerHisArgProLysAlaAlaAspMetAspLeuGluTyrArgGlnGlnGlySerGlyAlaArg 1562  
 Db 4523 TCTGCTGGCCAGGCGACAGCGTGTGCTGGAGTGGCGTCCCGGCTCCACAGCGCAG 4582  
 QY 1563 MetIleLeuGlnAspGluAspIleThrLysIleGluAsnAspTrpLysArgLeuAsn 1582  
 Db 4583 ---ATFCTGTGCGAGCTGGAGCTGACGTACAGCGGAGCGCGGTGGAGCGGCTCAAC 4639  
 QY 1583 ThrLeuAlaHisTyrGlnValProAspGlySerValValAlaLeuValSerLysGlnVal 1602  
 Db 4640 ACCCTTATCCACTACAATGTCCGGATGAGCCACCTCATCTCTG--- 4684  
 QY 1603 ThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSerAlaSerLysTyrGlu 1622  
 Db 4685 -----TCCAAAGTGGGGTCTCCACGACGCGGAGGACGCCAGCAG 4726  
 QY 1623 AsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgThrProMetIleThr 1642  
 Db 4727 GACCTG-----CCTGGGAGCGCCATGCCCTC--- 4753  
 QY 1643 ProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHisGluHisGlyAsp 1662  
 Db 4754 -----CTGGAGGAGGAGAACCGGTGTGCACCTGTGCGCGGACGAGGTGGAC 4807  
 QY 1663 GlnLysGluGlyAspArgGlySer-----LysMetValSerGlu 1675  
 Db 4808 GAGGCAAGTCCAAAGAGAGCGCCTGTAAGAGAGGAGCGGAGGAGCCATCACCAG 4867  
 QY 1676 IleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspLeu 1695  
 Db 4868 ATCTACTGACGCGGTCTCTCAGTCAAGGCGACACTGCGAGGAGTTGTGGACAATTC 4927  
 QY 1696 PheGluThrIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyr 1715  
 Db 4928 TTTCCAGAGGTGTGTCGCGCTGGGAC-----CGGTGGCCACCTGACGTCAAGTAC 4978  
 QY 1716 MetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArg 1735  
 Db 4979 TTTCTTCGACTCTCTGGAGCAGCAGGAGAGAGCAACATCCAGGATGAAGAGACCATC 5038  
 QY 1736 HisThrTrpLysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnPro 1755  
 Db 5039 CACATCTGGAAGACGACAGCTTACCCTCGGTCTGGGTGAACATCCTCAAGAACCCTC 5098  
 QY 1756 GlnPheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValAla 1775  
 Db 5099 CACTTCATCTTTCAGCTGATGTCACGAGGTGTGGAGCGCTCGCTGTCAGTCAATCGC 5158  
 QY 1776 GlnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSer 1795  
 Db 5159 CAGACCTTCATGATGCTGTCACGCGCAGGAGATAGCTGAGCGGATTCCTCCAGC 5218  
 QY 1796 AsnLysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyr 1815  
 Db 5219 AACAAAGTGTGTACGCGCAGGAGATCTCCACCTACAAGAGATGCTGGAGGATTACTAC 5278  
 QY 1816 SerAspIleGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrTrpLeuAlaGlu 1835  
 Db 5279 AAGGGATCCGCGAGATGTCAGGTACAGCAGGAGGACATGAACACACACCTGCGAG 5338  
 QY 1836 GlnSerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSer 1855  
 Db 5339 ATTTCCCGGCGCACACGCGACTCTCTGAACACCTCTGTCGGCCTCCACCTGCTACCAA 5398  
 QY 1856 TyrValGlyLysTyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCysGly 1875  
 Db 5399 TACACGGAAGATCTATGACGAGATCATCAATGCTTGGAGGAGATCTCTGCGGCCAG 5458



182	Lys	Leu	Phe	Ile	Ala	Thr	Ala	Val	Asp	Gly	Lys	Pro	Glu	Tyr	Phe	Pro	Thr	Ile	Ser	201
	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	:::::	
506	CGG	TCT	GCT	GTT	GTG	GGCA	AAAG	CAAG	TGG	CG	CAC	ACG	CA	CA	CG	CA	CG	CA	CT	565
	202	Arg	Lys	Leu	Thr	Lys	Asn	Ser	Glu	Ala	Asp	Gly	Met	Phe	Ala	Tyr	Val	Phe	His	221
566	CGG	CTG	TGG	ACG	GACT	CAC	AGC	AGG	GAG	CGC	TTG	AAG	CG	CT	TAC	ACG	CA	CG	CA	625
	222	Phe	Val	Ala	Ser	Met	Ile	Lys	Ile	Pro	Ser	Asp	Thr	Phe	Thr	Ile	Leu	Pro	Asp	241
626	TAC	AAG	CGC	GGC	TAC	CTG	TCC	ACC	AA	CAC	AC	AG	CAG	CAG	TTC					664
	242	Ile	Tyr	Tyr	Val	Tyr	Gly	Phe	Ser	Ser	Gly	Asn	Phe	Val	Tyr	Phe	Leu	Thr	Leu	261
665	-----	GTG	CGC	GGC	CTT	CG	AGC	AGC	CGC	CGC	CT	ACG	TCT	CTT	CTT	CTT	CTT	CTT	CTT	715
	262	Glu	Met	Val	Ser	Pro	Gly	Ser	Thr	Thr	Lys	Glu	Gln	Val	Tyr	Thr	Ser	Lys	Leu	281
716	GAC	AAG	--	-	CAC	CGC	GGC	CGG	AAC	CGC	CAG	C								751
	282	Arg	Leu	Cys	Lys	Glu	Asp	Thr	Ala	Phe	Asn	Ser	Tyr	Val	Glu	Val	Pro	Ile	Gly	301
752	CGA	TG	TGC	AGA	AAG	AC	CC	CA	CT	ACT	CT	CT	CT	CT	CT	CT	CT	CT	CT	811
	302	Arg	Ser	Gly	Val	Glu	Tyr	Arg	Leu	Leu	Gln	Ala	Ala	Tyr	Leu	Ser	Lys	Ala	Gly	321
812	GAC	CCC	GCA	TAC	CCAC															847
	322	Leu	Gly	Arg	Thr	Leu	Gly	Val	His	Pro	Asp	Asp	Ala	Leu	Phe	Thr	Val	Phe	Ser	341
848	CTG	CGC	CGC	TCC	GTG	CGC	TG	CT	CT	CGC	AGG	TG	CT	AT	TAT	CT	CT	CT	CT	907
	342	Gly	Gln	Lys	Arg	Lys	Met	Lys	Ser	Leu	Asp	Glu	Ser	Ala	Leu	Cys	Ile	Phe	Ile	361
908	GAC	AGC	-----	-	CGG	AGC	ACT	GGG	GGC	CGG	TG	CGG	CGC	CT	CT	CGC	TG	TCC	CGC	961
	362	Gln	Ile	Asn	Asp	Arg	Ile	Lys	Glu	Arg	Leu	Gln	Ser	Cys	Tyr	Arg	Gly	Glu	Gly	391
962	AAG	GTG	CAC	CG																

[illegible]

Db 2459 GGGGGCATCCGATCACCATTCTGGGTCCTCAATTTGGGGCTCCAAAGCAGGGGACATC--- 2515  
QY 892 SerHisValLysValAlaGlyValGluCysSerProLeuValAspGlyTyrIleProAla 911  
Db 2516 CAGAGATCTCTGTGGCGCGGGAAGTCTCTTTACGCGGAAAGCTTACTCCGTGTC 2575  
QY 912 GluGlnIleValCysGluMetGlyGluAlaLysProSerGlnHisAlaGlyPheValGlu 931  
Db 2576 ACCCGATCGTGTGTGATC---GAGGCTGGGAGACGCTTTCACGGGGGTGTCGAG 2632  
QY 932 IleCysValAlaValCysArgProGluPhe-----MetAlaArgSerSer----- 946  
Db 2633 GTGGAGTC-----TTGCGGAACCTGGCGGTTTCGCTCCCAATGTC 2674  
QY 947 GlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMet 966  
Db 2675 CAGTTCACCTTCCAACAGCCCAAGCCTCTCAGT---GTGGAGCCGAGCAGGACCGCAG 2731  
QY 967 SerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnVal 986  
Db 2732 GCGGGCGGCACACACTGACCATCCACGCCACCTGGACACGGCTCCAGGAGGAC 2791  
QY 987 ValMetPheGlyLysGlnProCysLeuPheHisArgSerProSerTyrIleValCys 1006  
Db 2792 GTG-----CGGGTGACCTCAACCGGCTCCCGTGT 2821  
QY 1007 AsnThrThrSerAspGluValLeuGluMetLysValSerValGlnValAspArgAla 1026  
Db 2822 AAGTCACCAAGTTTGGGCGGAGCTCCAGTGTCTACTGGCCCCCAGGGCAGCGGGC 2881  
QY 1027 LysIle-----HisGlnAspLeuValPhe 1034  
Db 2882 CAGATCTTCTGGAGGTCTCTACGGGGGTCCCGGTGCCCAACCCCGCATCTTCTTC 2941  
QY 1035 GlnTyrValGluAspProThrIleValArgIleGluProGluTrpSerIleValSerGly 1054  
Db 2942 ACTACCGGAAACCCCTACTGCGAGGCTTCGAGCCGTACGAAGCTTTGCCAGTGT 3001  
QY 1055 AsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArg 1074  
Db 3002 GCGCGCAGCATCAACGTCACGGGTCCAGGCTTCAGCCTGATCCAGAGTTTGCATG--- 3058  
QY 1075 AlaLysHisGlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMet 1094  
Db 3059 -----GTGGTCATCGCGGAGCCCTGCGAGTCTCGGCACGCG 3094  
QY 1095 ThrCysGlnAla-----ProAlaLeuAlaLeuGlyProAspHis----- 1107  
Db 3095 CCGGGGAGGCTGAATCCCTGCGAGCCCATGACGGTGTGGGTACAGCTACGTGTTCCAC 3154  
QY 1108 -----GlnSerAspLeuThrGluArgProGluGluPheGly 1119  
Db 3155 AATGACACCAAGTCGTCTTCTGTCTCCCGGTGTGCTGAGGAGCAGAGGCTTACAAC 3214  
QY 1120 PheIle-----LeuAspAsnValGlnSerLeuIleLeuAsnLysThrAsn 1135  
Db 3215 CTCACGGTCTGATCGAGATGGACGGCACCGTGCCTGTCTGAGAAGCAGCGCGGCC 3274  
QY 1136 PheThrTyrTyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeu 1155  
Db 3275 TTCGAGTACGTGCTGACCCCTTTGAGAACTTC----- 3310  
QY 1156 LysProGlyThrProIleIleLysGlyLysAsnLeuIleProProValAlaGlyGly 1175  
Db 3311 -----ACAGGTGGCTCAAGAAGCAGGTCAACAAGCTCATCCACGCCCGGGGCACC 3361  
QY 1176 AsnValLysLeuAsnTyrThrVal-----LeuValGlyGluLysProCys 1190  
Db 3362 AATCTGAACAGGCGATGACGTGTCAGAGGCGCGAGGCGCTTCTGTGGTGGCGGCGCTGC 3421  
QY 1191 ThrVal---ThrValSerAspValGlnLeuLeuCysLysSerProAsnLeuIleGlyArg 1209  
Db 3422 ACATGAAGACGCTGACGGAGACCGACCTGTACTGTGAGCCCCCGGAGGTGACGCCCCG 3481

QY 1210 HisLys-----ValMetAlaArgVal 1216  
Db 3482 CCCAAGCGCGCAGAAACGAGACACACACACCACTGCCCGAGTTTCATTGTGAGTTTC 3541  
QY 1217 GlyGlyMetGluTyrSerProGlyMetValTyrIle-----AlaProAspSerPro 1233  
Db 3542 GGCCTCTCGAGTGGTGTGGCGCGCTGGAGTACGACACACGCGGTGAGCAGCGTCCG 3601  
QY 1234 LeuSerLeu-----ProAlaIleValSerIleAlaValAlaGly 1246  
Db 3602 CTCAGCTCATCTTTCGCCCTGGTTCATCGTCCCATGGTGGTGTCTCATCGCGTGTCT--- 3658  
QY 1247 GlyLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrLysArgLysSerArgGlu 1266  
Db 3659 -----GTCTACTGCTACTGGAGGAGACCCAGCAG 3688  
QY 1267 SerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAla 1286  
Db 3689 GCCGAACGAGAGTATGAGAAGATCAAGTCCAGCTGGAGGGCTGGAGGAGCGTGGCG 3748  
QY 1287 LeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIleHisGluLeuThrSerAsp 1306  
Db 3749 GACCGCTGCAAGAAGAAATTCACAGACCTGATCGAGATGGAGGAGACCAACAC 3808  
QY 1307 LeuAspGlyAlaGlyIleProPheLeuAspTyrArgThrTyrMetArgValLeuPhe 1326  
Db 3809 GTGCACGNGCGCGATCCCGTGTGACTACAAGACCTACACGACCGCTCTCTCTTC 3868  
QY 1327 -----ProGlyIleGluAspHisProValLeuArgAspLeuGluValProGly 1342  
Db 3869 CTGCCCTCCAAAGGACGCGACAGACGTGATGATCACGGCAAGCTGGACATCCCTGAG 3928  
QY 1343 TyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGluLeuIleAsnLys 1362  
Db 3929 CCGCGCGCGCGGTGGTGGAGCGCCCTCTACCAAGTTCTCCAACTGCTGACACCAAG 3988  
QY 1363 ValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPheSerMetArgAsp 1382  
Db 3989 TCTTTCTCTCATTAATTCATCCACACCTGGAGAACAGCGGGAGTTCTCGCGCGCGCC 4048  
QY 1383 ArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeuGluTyrAlaThr 1402  
Db 4049 AAGTCTACTTTCGGGTCTCTGACGTGGCGCTGCACGGGAAACTGGAGTACTACAG 4108  
QY 1403 AspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuLysLysAsnHis 1422  
Db 4109 GACATCATGACACGCTCTTCTGAGGCTCTGGAGCAGTACGTGGTGGCCCAAGAAC--- 4165  
QY 1423 ProLysLeuLeuLeuArgArgThrGluSerValAlaGluLysMetLeuThrAsnTrpPhe 1442  
Db 4166 CCCAAGCTGATGCTGCGCAGGCTCGAGACTGTGGTGGAGAGGATGCTGTCCAACCTGGATG 4225  
QY 1443 ThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeuPheSerLeuPhe 1462  
Db 4226 TCCATCTCCCTGTACCACTAGCTCAAGACAGTGGCGGGAGCCCTGTACAAGCTTTC 4285  
QY 1463 CysAlaIleLysGlnGlnMetLysGlyProIleAspAlaIleThrGlyGluAlaArg 1482  
Db 4286 AAGCCATCAACATCAGGTGAAAGGCGCGGTGGATGGCGGTACAGAAAGGCCCAAG 4345  
QY 1483 TyrSerLeuSerGluAspLysLysIleArgGlnGlnIleAspTyrLysThrLeuValLeu 1502  
Db 4346 TACATCTCAACGACACCGGGCTGCTGGGGATGATGTGGAGTACGCCACCCCTGACGGTG 4405  
QY 1503 SerCysValSerProAspAsnAlaAsnSerProGluValProValLysIleLeuAsnCys 1522  
Db 4406 AGCGTGTGTCGAGGAGCAG---GGAGTGGACGCCATCCCGGTGAAGGTCTCAACTGT 4462  
QY 1523 AspThrIleThrGlnValLysLysIleLeuAspAlaIlePheLysAsnValProCys 1542  
Db 4463 GACACCATCTCCAGGTCAAGGAGAAAGATCATTTGACAGGTTGATCCGCGGCGACCCCTGC 4522



CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a cDNA encoding a  
 CC novel human polypeptide of the invention.

XX  
 SQ Sequence 6329 BP; 1077 A; 1939 C; 2034 G; 1279 T; 0 other;

## Alignment Scores:

Pred. No.: 3,62e-199 Length: 6329  
 Score: 2539.00 Matches: 652  
 Percent Similarity: 50.43% Conservative: 343  
 Best Local Similarity: 33.05% Mismatches: 716  
 Query Match: 25.42% Indels: 262  
 DB: 22 Gaps: 61

US-09-964-956-13 (1-1896) x ABA09062 (1-6329)

QY 24 ThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPhe 43  
 DB 6232 ACCCTGCTGGCGTGTGGCGCAGGTGCCAGCTGAGCGCCCGCAAGCTGGAGCTTCTTC 6173  
 QY 44 ArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIle 63  
 DB 6172 CGCAGCGAG-----AAAGAGCTGAGACCACTCGCTGGTGTGATGAGCGCTCAGCGGTGGTG 6119  
 QY 64 TyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThr 83  
 DB 6118 TACCTGGGCGGTGAATGCCCTCTACAGCTGGATGCGAAGCTGCAGCTGGAGCAGCAG 6059  
 QY 84 HisGluThrGlyProAspGluAspAsnProLysCystyrProProArgIleValGlnThr 103  
 DB 6058 GTGGCCAGCGGCGGCTCTGGAACAACAAGAGTGCACGCGCCCATCGAGCGGCAGCAG 5999  
 QY 104 CysAsnGluProLeuThrThrAsnAsnValAsnLysMetLeuLeuLeuLeuLeuLeuLys 123  
 DB 5998 TGGCATGAG---GCTAGATGACTGACATGTCAACAGCTGTGCTGGTGCAGCCCTCC 5942  
 QY 124 GluAsnArgLeuLeuAlaCysGlySerLeuTyrGlnGlyIle-CysLysLeuLeuArgLe 143  
 DB 5941 AGAAACAGCGCTGTGTGAGTGGCGGAGCTTCTTAAGGGCATCTCGCTCTGCGCGCCT 5882  
 QY 143 uGluAsp-----LeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeuSe 161  
 DB 5881 GAGCAACATCTCCCTCGGCTGTCTTACGAGGAGCGCAGCGGGGAGAAAGTCTTTCGTGGC 5822  
 QY 161 rGlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspAs 181  
 DB 5821 CAGCATGATGAGGGCGTGGCCACAGTGGGCTGGTGGTGCAGCTCCAGGGTCTCGTGGTGA 5762  
 QY 181 pLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSe 201  
 DB 5761 CGCGCTGTGTGTGGGAAAGGCAATGGGCCACACAGCAACGCGCATCATCTGTGAGCAC 5702  
 QY 201 rArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisaspG 221  
 DB 5701 TGGGCTGTGGAGGAGTACAGCAGGAGGAGCGCTTTGAAGCCCTACAGCAGCAGCAGCAC 5642  
 QY 221 uPheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAs 241  
 DB 5641 CTACAGGCGGCTACTGTCCACCAACACACAGCAGCTC----- 5602  
 QY 241 pIleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnPr 261  
 DB 5601 -----GTGGCGGCTTCGAGGAGGCGCCCTACGCTCTTCTTCTTCAACACGACA 5552  
 QY 261 oGluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuVa 281

DB 5551 GGACAAG---CACCGCGCCCGGAACCGCAGC-----CTGCTGCG 5516  
 QY 281 LArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCysG 301  
 DB 5515 ACGGATGCGAGAGAAAGACCCCACTACTACTCTACCTGGAGATGGACCTGACGTGCGG 5456  
 QY 301 uArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaVa 321  
 DB 5455 GGACCCCGACATCCAC-----GCCGCTGCCCTTT-----GGCACCTG 5420  
 QY 321 LLeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerLys 341  
 DB 5419 CCTGGCGGCTCCGCTGGCTGGCTGCTGCTGGCAGGTGTATATGCTGTCTTACGACG 5360  
 QY 341 sGlyGlnLysArgLysMetLysSerLeuAspLysSerAlaLeuCysIlePheIleLeu 361  
 DB 5359 AGACAGC-----CGGACAGTGGGGCGCGGTGGCGGCTCTGCTGCTTCCCGCTGA 5306  
 QY 361 sGlnIleAsnAspArgIleLysGluArgLeuGlnSerCystyrArgGlyGluGlyThrLe 381  
 DB 5305 CGAGGTGCACGCCCAAGATGGAGGCCAACCGCAACGCTGTTAC-----ACAGGCACCG 5252  
 QY 381 uAspLeuAlaTrpLeuLysValLys-----AspIleProCysSerSerAlaLe 397  
 DB 5251 GGAGGCGGCTGACATCTTCTACAAGCCCTTCCACGGCATATCCAGTGGCGGCGCACGC 5192  
 QY 397 uLeuThrIleAspAspAsnPhe---CysGlyLeuAsp---MetAsnAlaProLeuGlyVa 415  
 DB 5191 GCCGGGCTCCAGCAGAGCTTCCCATGCTGGCTGGAGCACCTGCCCTACCCCTGGGCGAG 5132  
 QY 415 LserAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThrSerVa 435  
 DB 5131 CGCGCGGGCTCAGAGGCACACGCGTCTGCAGCGTGGAGCGCTGAACCTCACGCCCT 5072  
 QY 435 LileAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLe 455  
 DB 5071 GACGTCGCCCGCGGAACACACACTGTGTCTTCTGGGCACCTCTGATGGCGGAT 5012  
 QY 455 uLysIleArgValAspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGlnVa 475  
 DB 5011 CCTCAGGTGATCTCAC---CCAGATGGGACCTCTCCAGAGTACGACTCTATCTTGT 4955  
 QY 475 LValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuTy 495  
 DB 4254 GCAGATAAACAAGAGAGTCAAGCGCGACTGTGTCTGTGGAGCCTGGCGGACGCTGTA 4895  
 QY 495 rIleMetSerGluArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrGlnSe 515  
 DB 4894 CGCATGACCCAGGACAAGGTTCCTCGGCTGCCGTGGAGGAGTGCCTGAGCTACCCGAC 4835  
 QY 515 rCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTyrCysValLeuHisAsnPh 535  
 DB 4834 CTGCACCGAGTCCCGGACTCCAGGACCTCTACTGCGGCTGTGCTGCTGCGAGGAGC 4775  
 QY 535 rCysThrArgLysGluArgCysGluArgSerLysGluProArgArgPheAla---SerG 554  
 DB 4774 ATGCACCGGAGGCGGAGTGTCCGCGGCGGAGGAGCCACCTGCTGTGGAGCGG 4715  
 QY 554 uMetLysGlnCysValArgLeuThr---ValHisProAsnAsnIleSer---ValSerG 572  
 DB 4714 AAGCAAGTCTCGCTGGCGGCTCCAGCGGCCCGCCAGCAACATGAGCGCGGCGGCGCA 4655  
 QY 572 nTyrAsnValLeuLeuValLeuThrTyrAsnValProGluLeuLeuSerAla-----G 590  
 DB 4654 GGGGAGTGCAGCTGACCGCTCAGCCCC-----CTCCCTGCCCTGAGCGAGGAGCACA 4601  
 QY 590 yValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValGlyAsnGlnI 610  
 DB 4600 GTTGTGCTGCTCTTTTGGGAGTCCGCCACACCCCGCGCTGGAGGCGGAGCGGCT 4541  
 QY 610 eGlnCysTyrSerProAlaAlaLysGluValProArgIleThrGluAsnGly---As 629

Db 4540 CATCTGCACTCCCAAGCAGC-----ATCCCC-----GTCACACCGCCAGGCCAGGA 4493  
Qy 629 pHisHisValValGlnLeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerTh 649  
Db 4492 CCACGTGGCGGTACCATCCAGCTCCCTTAGACGAGGCAACATCTTCCTCAGCTCCTA 4433  
Qy 649 rSerPheValPheThrAsnCysSerValHisAsnSer-----CysLe 663  
Db 4432 CCAGTACCCCTTTACAGCTCCGCCAGGCCATGAGCTGGAGGAGAACCTGCCGTGCAT 4373  
Qy 663 userCysValGluSerProTyrArgCysHisTrpCysLysTyrArgHisValCysThrH1 683  
Db 4372 CTCTCGGTGAGCAACCGCTGGACCTGCCAGTGGACCTACACAGAGTCCGGGA 4313  
Qy 683 s---AspProLysThrCysSerPheGlnGluCly-----ArgValLysLeuProGluAs 700  
Db 4312 GGCTTCGCCCAACCT-----GAGGAGGGATCGTCGTGCCCATATGGAGGACAG 4262  
Qy 700 pCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysProI1 720  
Db 4261 CTGTCCTCCAGTCTCTGGGACCCAGCCCTCGTATCCCATGAACACACGACAGATGT 4202  
Qy 720 eThrLeuLysAlaLysAsnLeuProGlnSerGlyGlnArgGlyTyrGluCysI1 740  
Db 4201 GAACCTCCAGGGCAAGAACCTGGAC----- 4177  
Qy 740 eLeuAsnIleGlnGlySerGluGlnArgValProAla-----LeuArgPheAsnSerSe 758  
Db 4176 ---ACCGTGAAGGTTCTCTCCCTGACGTGGGACGTACTGTCTCAAGTTATCGAGCC 4121  
Qy 758 rSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsn----- 776  
Db 4120 GGTGACCATGCAGGAATCTGGACCTTCGCTTCGGACCCCAAGCTGTCCACGATGC 4061  
Qy 777 -----LeuProValGluLeuThrValValTrpAsnGlyHisPheAsnIleAspAs 793  
Db 4060 CAACGAGACGCTGCCCTTCACCTTTACGTCAAGTCTTACGGCAAG---AATATCGAC--- 4006  
Qy 793 nProAlaGlnAsnLysValHis-----LeuTyrLysCysGlyAlaMetArgGluSerCy 811  
Db 4005 -----AGCAAGCTCCATGTGACCTCTACGACTGCTCTTTGGCCGCGAGGACTG 3956  
Qy 811 sGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTrpCysGlnGlyProGlyG1 831  
Db 3955 CAGCTGTGCGGGCGCGCTAAACCCGACTACAGGTGTGCGTGTGCGGGGCGCAGAGCAG 3896  
Qy 831 nCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuGluLeuSerGlyAl 851  
Db 3895 GTGCTGTATAGGCCCTGTGC-----AACAC 3869  
Qy 851 aLysSerLysCysThrAsnProArgIleThrGluIleProValThrGlyProArgG1 871  
Db 3868 CACCTCCGAGTCCCGCGCGCTCATCACCAGATCCAGCTGAGAGCGGCCCTGGG 3809  
Qy 871 uGlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPheArgAspIleAl 891  
Db 3808 TGGGGGCATCCGATCACCATCTCGGGTCCAATTGGCGTCCCAAGAGGGGACATC--- 3751  
Qy 891 aserHisValLysValAlaGlyValGluCysSerProLeuValAspGlyTyrIleProAl 911  
Db 3750 -CAGAGGATCTCTGCGCGCGCGGAGTCTCTTACCGGGAACGTACTCCGTGTC 3692  
Qy 911 aGluGlnIleValCysGluMetGlyGluAlaLysProSerGlnHisAlaGlyPheValG1 931  
Db 3691 CACCCGGATCGTGTGTGATC---GAGGCTCGGAGACGCTTTTCACGGGGGGTGTCTGA 3635  
Qy 931 uileCysValAlaValCysArgProGluPhe-----MetAlaArgSerSer----- 946  
Db 3634 GTGGACGTC-----TTCGGGAACCTGGCGGCTTCGCTCCCAATGT 3593  
Qy 947 -GlnLeuTyrTyPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMe 966  
Db 3592 CCAGTTCACCTTCCACACGCCCAAGCTCTCACT---GTGAGCGCGCAGGAGGCCGCA 3536

Qy 966 tSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnValVa 986  
Db 3535 GCGGGCGGACCACTGACCATCCACGCGACCCACCTGGACACGCGGTCCCGAGGAGGA 3476  
Qy 986 lValMetPheGlyLysGlnProCysLeuPheHisArgArgSerProSerTyrIleValCy 1006  
Db 3475 CGTG-----CGGTGACCCCTCAACGCGCTCCGCTG 3446  
Qy 1006 sAsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnValAspArgAl 1026  
Db 3445 TAAAGTGACGAAGTTTGGGCGCAGCTCCAGTGTGTCACTGGCCCCCAGCGACACGGGG 3386  
Qy 1026 alysile-----HisGlnAspLeuValPh 1034  
Db 3385 CCAGATGCTTCTGGAGTCTCTACGGGGGTCCCGCTGCCCAACCCCGGCATCTCTT 3326  
Qy 1034 eGlnTyrValGluAspProThrIleValArgIleGluProGluTrpSerIleValSerG1 1054  
Db 3325 CACCTACCGCGAAACCCCGTACTCGAGCGCTTCGAGCGCTACGAGAGCTTTCGCCAGTGG 3266  
Qy 1054 yAsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleAr 1074  
Db 3265 TGGCCCGCACCATCAAGCTCAGCGTTCAGGCTTCAGCTGATCCAGAGGTTTGCATG--- 3208  
Qy 1074 gAlaLysHisGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMe 1094  
Db 3207 -----GTGGTCACTCGGGAGCCCTCGAGTCTCGAGTCTGCGAGCC 3173  
Qy 1094 tThrCysGlnAla-----ProAlaLeuAlaLeuGlyProAspHis----- 1107  
Db 3172 GCGCGGGAGGCTGAATCCTCGAGCCCATGACGCGTGGTGGGTACAGCTACGTGTCCA 3113  
Qy 1108 -----GlnSerAspLeuThrGluArgProGluGluPheG1 1119  
Db 3112 CAATGACACCAAGTCTCTCTCCTGTCGCCGGGTGCTGCTGAGGAGCAGAGGCTACAA 3053  
Qy 1119 yPheIle-----LeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrAs 1135  
Db 3052 CCTCAGCGTCTGATCGAGATGGAGCGGACCGCTGCTCAGACACAGAGCGCGGGC 2993  
Qy 1135 nPheThrTyrTyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLe 1155  
Db 2992 CTTCGAGTACGTCCTCGACCCACCTTTGAGAACTTC----- 2956  
Qy 1155 uLysProGlyThrProIleIleLeuLysGlyLysAsnLeuIleProProValAlaGlyG1 1175  
Db 2955 -----ACAGGTGGCTCAAGAAGCAGGTCAACAAGCTCATCCCGCGCCCGGGGCAC 2906  
Qy 1175 yAsnValLysLeuAsnTyrThrVal-----LeuValGlyGluLysProCy 1190  
Db 2905 CAATCTGACACAGCGCATGACGTGAGGAGGCGCGAGCGCTTCTGGTGGCGGCGCTG 2846  
Qy 1190 sThrVal---ThrValSerAspValGlnLeuLeuCysGluSerProAsnLeuIleGlyAr 1209  
Db 2845 CACCATGAGACGCTGAGCGAGACCGACCTGTACTGTGAGCCCGCGAGGTGACGCCCC 2786  
Qy 1209 gHisLys-----ValMetAlaArgVa 1216  
Db 2785 GCCCAAGCGCGCAGAAACAGACACACACACACCTGCGCGAGTTCAATTGTGAAGTT 2726  
Qy 1216 lGlyGlyMetGluTyrSerProGlyMetValTyrIle-----AlaProAspSerPr 1233  
Db 2725 CGGCTCGCGAGTGGTGTGGCCCGGTGGAGTACACACACCGGTGAGCGAGCTGCC 2666  
Qy 1233 oLeuSerLeu-----ProAlaIleValSerIleAlaValAlaG1 1246  
Db 2665 GCTCAGCTCATCTTTCGCTGGTGTCTGTCCTGCTGCTCGCGGTGCTC--- 2608  
Qy 1246 yGlyLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrLysArgLysSerArgG1 1266  
Db 2607 -----GTCTACTGCTACTGGAGAGAGCCAGCA 2579





4834 CTGCACCCAGTCGCGGACTCCAGAGACCCCTACTCTCGGCTGGTGGTCTCGTCCAGGAGC 4775  
535 rCysThrArgLysGluArgCysGluArgSerLysGluProArgPheAla---SerGI 554  
4774 ATGCACCCGGAAGCGGAGTGTCCGCGGCGGAGGAGCCAGCCACTGCTGTGGAGCG 4715  
554 uMetLysGlnCysValArgLeuThr---ValHisProAsnAsnIleSer---ValSerGI 572  
4714 AAGCAAGTCTCGTGGCGCTCACCAGCGCCAGCCAGCCAGACATAGCGCGCGGCCCA 4655  
572 nTyrAsnValLeuValLeuGluThrTyrAsnValProGluLeuSerAla-----GI 590  
4654 GGGGAGGTGCAGTGCAGCTCAGCCCC-----CTCCCTGCCTGAGCGAGGAGCAGA 4601  
590 yValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValcylAsnGlnII 610  
4600 GTTGTGTGCCCTTTTGGGAGTGGCGGCCAGCCAGCCCGCGCGGTGGAGGCGGCGGT 4541  
610 eGlnCysTyrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGly---As 629  
4540 CATCTGCAACTCCCAAGCAGC-----ATGCC-----GTACACCCGCCAGGCCAGGA 4493  
629 pHisHisValValGlnLeuGlnLysSerLysGluThrGlyMetThrPheAlaSerTh 649  
4492 CCACGTGGCGGTGACCATCTCAGTCTCTCTAGACGAGGCAACATCTCTCTCAGTCTCA 4433  
649 rSerPheValPheTyrAsnCysSerValHisAsnSer-----CysLe 663  
4432 CCAGTACCCTCTTACGACTGCGCGCAGCCATGAGCCTGGAGGAGAACCTGCGGTGCAT 4373  
663 uSerCysValGluSerProTyrArgCysHisTrpCysLysTyrArgHisValCysThrHi 683  
4372 CTCCTCGGTGACCAACCGCTGGACCTCGCAGTGGGACCTCGCTACCCAGCAGTGGCGGA 4313  
683 s---AspProLysThrCysSerPheGlnGluGly-----ArgValLysLeuProGluAs 700  
4312 GCCTTCGCCCAACCCCT-----GAGGAGGCATCGCTCGCGCCACATGAGGAGCAG 4262  
700 pCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysProII 720  
4261 CTGTCGCCAGTCTCTGGGACCCAGCCCTGGTGTATCCCTCCATCAACACGAGACAGATG 4202  
720 eThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCysII 740  
4201 GAACCTCCAGGCGCAGAACCTGGAC-----GAGGAGGAGGAGGAGGAGGAGGAGGAG 4177  
740 eLeuAsnIleGlnGlySerGluGlnArgValProAla-----LeuArgPheAsnSerSe 758  
4176 ---ACCCTGAGGGTCTCTCCCTGCGAGTGGCAGTACTGCTCAAGTTCTATGGAGCC 4121  
758 rSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsn----- 776  
4120 GTGACCATGACAGGATCTGGGACCTTCGCCTTCGGACCCCAAGCTGCCACCATGC 4061  
777 -----LeuProValGluLeuThrValValTrpAsnGlyHisPheAsnIleAspAs 793  
4060 CAACGAGACGCTGCCCTGACCTTTACGTCAAGTCTTACGGCAAG---AATATCGAC-- 4006  
793 nProAlaGlnAsnLysValHis-----LeuTyrLysCysGlyAlaMetArgGluSerCy 811  
4005 -----ACCAAGTCCATGTGACCTCTACGACTGCTCTTTGGCGGAGGAGGAGGAGGAG 3956  
811 sGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTyrPcysGlnGlyProGlyGI 831  
3955 CAGCCTGTGCCGGCGCGCTAAACCCGACTACAGTGTGGTGGTGGCGGGGCCAGAGCAG 3896  
831 nCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuLeuLeuSerGlyAl 851  
3895 GTGCGGTATGAGGCCCTGTGC-----AACAC 3869  
851 aLysSerLysCysThrAsnProArgIleThrGluIleIleProValThrGlyProArgGI 871  
3868 CACCTCCGAGTGCCTCCCGCGCTCATACACGAGGATCCAGCCTGACAGCGGCCCTCGG 3809

871 uGlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPheArgAspIleAl 891  
3808 TGGGGCATCGCATCACCATCTGGGTCCAATTTGGCGCTCCAAGGAGGAGCATC-- 3751  
891 aSerHisValLysValAlaGlyValGluCysSerProLeuValAspGlyTyrIleProAl 911  
3750 -CAGAGGATCTCTGCGCGCGGAACTGCTCTTTCAGCGGAACTGTTACCGGGGTGTCA 3692  
911 aGluGlnIleValCysGluMetGlyGluAlaLysProSerGlnHisAlaGlyPheValGI 931  
3691 CACCGGATCGTGTGTGATC---GAGCTCGGAGAGCGCTTTCACCGGGGGTGTCA 3635  
931 uIleCysValAlaValCysArgProGluPhe-----MetAlaArgSerSer----- 946  
3634 GGTGAGCTC-----TTGGGAAACTGGGCGCTTGGCTCCCAATGT 3593  
947 -GlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMe 966  
3592 CCAGTTCACCTTCCAACGCCCAAGCTCTCAGT---GTGGAGCGCAGCAGGAGCCGA 3536  
966 tSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnValVa 986  
3535 GCGCGGGGACACACTGACCATCCAGCGCACCCACTGGACACGGCTCCACGAGGA 3476  
986 lValMetPheGlyLysGlnProCysLeuPheHisArgArgSerProSerTyrIleValCy 1006  
3475 CGTG-----CGGTGACCTCAACGGGTCCCGTC 3446  
1006 sAsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnValAspArgAl 1026  
3445 TAAAGTACGAAGTTTGGGCGCGACCTCCAGTGTGTCTACGTGCCCGCCAGCGGACAG 3386  
1026 aLysIle-----HisGlnAspLeuValPh 1034  
3385 CCAGATGCTTCTGGAGTCTCTACGGGGGTCTCCCGTGCCTCCACCGCGCATCTTCTT 3326  
1034 eGlnTyrValGluAspProThrIleValArgIleGluProGluTrpSerIleValSerGI 1054  
3325 CACCTACCGGAAACCCGCTACTCGAGCGCTTCGAGCGCTTACGAAGCTTTTGCAGTGG 3266  
1054 yAsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleAr 1074  
3265 TGGCCGAGCATCAACCTCAGCGGTTCAGCGCTTACGTGATCCAGAGGTTTGCATG-- 3208  
1074 gAlaLysHisGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMe 1094  
3207 -----GTGCTATCGCGGAGCCCTGCAGTCTCGCAGCC 3173  
1094 tThrCysGlnAla-----ProAlaLeuAlaLeuGlyProAspHis----- 1107  
3172 GCCCGGAGGCTGAATCCCTCGACGCCCATGCGGTGGTGGTACAGACTACGTGTTCAC 3113  
1108 -----GlnSerAspLeuThrGluArgProGluGluPheGI 1119  
3112 CAATGACACCAAGTCTCTCTCTGCTCCCGCTGCTGAGGAGCGAGGAGGAGGAGGAG 3053  
1119 yPheIle-----LeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrAs 1135  
3052 CCTCAGCGTGTGATCGAGATGGAGGCGCACCGTGCCTCTCAGACACAGAGCGCGGCG 2993  
1135 nPheThrTyrTrpProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLe 1155  
2992 CTTCAGTACCTGCTGACCCGCTTGTGAGAACTTC----- 2956  
1155 uLysProGlyThrProIleIleLeuLysGlyLysAsnLeuIleProProValAlaGlyCl 1175  
2955 -----ACAGTGGCGTCAGAGAGCTCAACAGCTCATCTCCCGCCCGCGGCGAC 2906  
1175 yAsnValLysLeuAsnTyrThrVal-----LeuValGlyGluLysProCy 1190  
2905 CAATCTGAACAGGCGGATGAGCTGCAGGAGGCGGCGGCTTCGTGGTGGCGGAGCGGTG 2846









Db 5017 GTGACCTTGAGTGGTTCGCTCCAGCACACAGAGCTACATCTTCGGGACCTGGAGC 5076  
 Qy 1571 ThrThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHisTyrGlnValPro 1590  
 Db 5077 ACCTCAGTGTGGGAAGCGCGCAAGAGCTTAACACGCTGGCCCATTAAGAATCCCT 5136  
 Qy 1591 AspGlySerValValAla-----LeuValSerLysGlnValThrAlaTyrAsnAlaVal 1608  
 Db 5137 GAAGTGCCCTCCCTGGCCATGAGTCTCATAGACAAG----- 5175  
 Qy 1609 AsnAsnSerThrValSerArgThrSerAlaSerLysTyrGluAsnMetIleArgTyrThr 1628  
 Db 5176 ---GACAACACACTGGCGCGAGTGA--- 5199  
 Qy 1629 GlySerProAspSerLeuArgSerArgThrProMetIleThrProAspLeuLysSerGly 1648  
 Db 5200 -----GACTTGGACACA--- 5211  
 Qy 1649 ValLysMetTrpHisLeuVal-----LysAsnHis 1658  
 Db 5212 GAGAAGTATTTCATTTGGTCTCCCTACGACGAGCTGGCGGAGCCCAAGAAGTCTCAC 5271  
 Qy 1659 GluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLeu 1678  
 Db 5272 CGGACAGCCATCGCAAG-----AAGGTGCTCCCGCAATCTACCTG 5313  
 Qy 1679 ThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspLeuPheGluThr 1698  
 Db 5314 ACCGCTGCTCTCCACAGGCGACGTTGAGAAAGTTTCTGGATGACCTGTTCAAGGCC 5373  
 Qy 1699 IlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAsp 1718  
 Db 5374 ATTCTGAGTATC-----CGTGAAGACAGCCCTGCTGCTCAAGTACTTTTTCGAC 5427  
 Qy 1719 PheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrp 1738  
 Db 5428 TTCCTGGAGGACGAGGTGAGAAAGGGGAAATCTCCGACCCCGACACCTACACATCTGG 5487  
 Qy 1739 LysSerAsnLysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheVal 1758  
 Db 5488 AAGACCAACAGCCTTCTCTCGGTTCGGGTGACATCTCTGAGAACCCCGCAGTTTGTG 5547  
 Qy 1759 PheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnThrPhe 1778  
 Db 5548 TTTGACATCGACAAGACAGACACATCGACGCTGCTTTCAGTTCATCGCGAGCCCTC 5607  
 Qy 1779 MetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeu 1798  
 Db 5608 ATCGACGCTGCTCCATCTCTGACCTGGCAAGGATCGCCCAACCAACAGCTC 5667  
 Qy 1799 LeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSerAspIle 1818  
 Db 5668 CTCTACGCCAAGGAGATTCCTGAGTACCGGAAGATCGTGACGCGTACTACAAGCAGATC 5727  
 Qy 1819 GlyLysMetProAlaIleSerAspClnAspMetAsnAlaTyrLeuAlaGluClnSerArg 1838  
 Db 5728 CAGGACATGACGCGCTCAGGACGAGAGATGATGCCCATCTGCGCGAGGAGTCGAGG 5787  
 Qy 1839 MethIleMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGly 1858  
 Db 5788 AATACCAGATGAGTTCACACCACTGTCGCCATGGCAGAGATTTAAGTACCCCAAG 5847  
 Qy 1859 LysTyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCysGlyLysGlnLys 1878  
 Db 5848 AGGTATGGCGCGAGATGCGCGCTGGAGGCCAACCCGACCGCCCGGAGGACACAA 5907  
 Qy 1879 LeuAlaTyrLysLeuGluClnValIleThrLeuMet 1890  
 Db 5908 CTGACGACACAAGTTTGACGAGGTGCTGCTTGATG 5943

RESULT 11

AAS00022

ID AAS00022 standard; cDNA; 5892 BP.

XX AAS00022;  
 AC 09-MAY-2001 (first entry)  
 DT Human cDNA encoding Plexin-D1.  
 DE Human; Plexin-D1; semaphorin domain; hyperplasia; neoplasia; cancer;  
 KW neurodegenerative disease; autoimmune disease; lupus; multiple sclerosis;  
 KW inflammatory bowel disease; diabetes type I; rheumatoid arthritis;  
 KW immunogen; antibody; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 CDS 13..5790  
 FT /\*tag= a  
 FT /\*product= "Plexin-D1"  
 FT 13..138  
 FT /\*tag= b  
 FT 139..5787  
 FT /\*tag= c  
 FT /\*label= "Mature Plexin-D1"  
 XX WO200114420-A2.  
 XX 01-MAR-2001.  
 XX 25-AUG-2000; 2000WO-US23365.  
 XX 25-AUG-1999; 99US-0150576.  
 XX (UYTO-) UNIV TORINO.  
 PA (REGC) UNIV CALIFORNIA.  
 XX Artigiani S, Comoglio PM, Goodman CS, Tesier-Lavigne M;  
 PI Tamagnone L;  
 DR WPI; 2001-226610/23.  
 P-PSDB; AAU00017.  
 XX New plexin polynucleotides and polypeptides, useful in diagnosis,  
 PT therapy and in producing compounds for treating diseases involving  
 PT aberrant cell growth (e.g. cancer) or immune regulation (e.g.  
 PT autoimmune diseases)  
 XX Claim 2; Page 62-64; 79pp; English.  
 XX The sequence encodes Human Plexin-D1. Plexins are large transmembrane  
 CC proteins whose extracellular domain shares homology with Scatter factor  
 CC receptors and contain an approximately 500 amino acid Semaphorin domain.  
 CC The plexin polynucleotides and polypeptides, and plexin-specific binding  
 CC agents are useful in diagnosis, therapy and in the biopharmaceutical  
 CC industry. In particular, the plexin polynucleotides and polypeptides are  
 CC useful for generating compounds (e.g. plexin-specific binding agents or  
 CC antibodies) for treating or diagnosing a disease or disorder involving  
 CC aberrant cell growth (e.g. hyperplasia, neoplasia, cancer or  
 CC neurodegenerative disease), or diseases or disorders involving aberrant  
 CC immune regulation (e.g. autoimmune diseases such as lupus, inflammatory  
 CC bowel disease or Diabetes Type I), or immunosuppressive diseases such as  
 CC multiple sclerosis or rheumatoid arthritis.  
 XX Sequence 5892 BP; 1121 A; 1976 C; 1767 G; 1028 T; 0 other;

## Alignment Scores:

Pred. No.: 5,74e-193 Length: 5892  
 Score: 2463.50 Matches: 649  
 Percent Similarity: 48.59% Conservative: 334  
 Best Local Similarity: 32.08% Mismatches: 750  
 Query Match: 24.66% Indels: 293  
 DB: Gaps: 59

US-09-964-956-13 (1-1896) x AAS00022 (1-5892)



Db 2101 TTCCACCTCTACGACTGCAGCCGCACTGCAACAGTGATACCCACACAGCGCTGTACACAGC 2160  
QY 665 CysValGluSerProTyrArgCysHisTrpCysLysTyrArgHisValCysThrHisAsp 684  
Db 2161 TGCCTTCGCGCAGCAGTGGCCCTTTCTGTGTGCGCAGCAGCAGCAGCTCTGTGTTCACAA 2220  
QY 685 ProLysThrCysSerPheGlnGluGlyArgValLysLeuProGluAspCysProGlnLeu 704  
Db 2221 CAGTCTCGGTGCGAGGCTCACCAACCCACAGCAGC---CCTCAGGAGTGCCTCCCGGACC 2277  
QY 705 LeuArgValAspLysLeuValProValGluValLysProLysThrLeuLysAla 724  
Db 2278 CTGCTCTACCCCTGCGCAGCCGCTTACCGGTGCTCCAGAACATCCTGGTGCCTCTG 2337  
QY 725 LysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyrGluCysLysLeuAsnIleGln 744  
Db 2338 GCCAACACTGCCCTTTTCCAGGGTGCA-----GCCCTGGAGTGTAGTTT----- 2382  
QY 745 GlySerGluGlnArgValProAlaLeuArgPheAsnSerSerValGlnCys---Gln 763  
Db 2383 GGCTGGAGGAGATCTTCGAGGCTGTGTGGTGAATGAGTGTGTGACGCTGTGACCA 2442  
QY 764 AsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsnLeuProValGluLeuThrVal 783  
Db 2443 GTGGTGTGCACAGCAGCCCGAAGAGCCAGGTG-----TTCCTCGCTCAGCCTCCAACTA 2496  
QY 784 ValTrpAsnGlyHisPheAsnIleAspAsnProAlaGlnAsnLysValHisLeuTyrLys 803  
Db 2497 AAGGGGGGGCAGCCCGATTCCTGGACAGCCCTGAGCCCATGACAGTCATGCTATTAAC 2556  
QY 804 CysGlyAlaMetArgGluSerCysGlyLeuCysLeu---LysAlaAspProAspPheAla 822  
Db 2557 TGTCCATGGCAGCCCGCAGTGTTCAGTGTCCAGTGTCCGGCCGCGAAGACCTGGGTCACTG 2616  
QY 823 CysGlyTrpCysGlnGlyProGlyGlnCysThrLeuArgGlnHisCysProAlaGlnGlu 842  
Db 2617 TGCCTGTGGAGTATGGC-----TGGCCGCTGCGGGG-----CCTCTGGAG--- 2658  
QY 843 SerGlnTrpLeuGluLeuSerGlyAlaLysSerLysCysThrAsnProArgIleThrGlu 862  
Db 2659 -----CCCATGGCTGGCACC-----TGCCCGCCCGCCGAGATCGCGCGC 2697  
QY 863 IleIleProValThrGlyProArgGluCysGlyThrLysValThrIleArgGlyGluAsn 882  
Db 2698 ATTGAGCCCTGAGTGGCCCTTGGAGGTGGGAGCCCTGACCATCGAGGAAGAAC 2757  
QY 883 LeuGlyLeuGluPheArgAspIleAlaSerHisValLysValAlaGlyValGluCysSer 902  
Db 2758 CTGGCCCGCGGCTCAGTGACGTGGCCACCGCGGTGGATGGTGGTGGCTGTGAG 2817  
QY 903 ProLeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMetGlyGluAlaLys 922  
Db 2818 CCACTGCTGCACATACACGTGTCGGAGGAGATCGTGTGTGCACAGGCGCAGCC--- 2874  
QY 923 ProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgProGluPheMet 942  
Db 2875 CCAGACCGCTCTCAGGTGTGTGACCGTG-----AACGCCCTCTAAGGAG 2919  
QY 943 AlaArgSerSerGlnLeuTyrPheMetThrLeuThrLeuSerAspLeuLysProSer 962  
Db 2920 GGCAAGTCCCGGACCGCTCTCTCTACGTGCTGCCCTGCTCCACTCCCTGGAGCCTACC 2979  
QY 963 ArgGlyProMetSerGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGly 982  
Db 2980 ATGGCCCAAGCCCGGGGACCAAGATCACCATCCATGGGAATCACCTCCATGTAGGC 3039  
QY 983 SerAsnValValMetPheGly---LysGlnProCysLeuPheHisArgArgSerPro 1001  
Db 3040 TCCGAGCTCCAGGCTCTGTGTGAACACACAGACAGCCCTGCAGCGAGCTGTATGCGACAGAT 3099  
QY 1002 SerTyrIleValCysAsnThrThrSerSerAspGluValLeuGluMetLysValSerVal 1021  
Db 1002 SerTyrIleValCysAsnThrThrSerSerAspGluValLeuGluMetLysValSerVal 1021

Db 3100 ACCAGCATCGCCTGC-----ACCATGCTCAGGGGGCCCTCCCGGCTCCCGTGCCTGTG 3153  
QY 1022 GlnVal-----AspArgAlaLysIleHisGlnAspLeuValPheGlnTyrValGlu 1038  
Db 3154 TGTGTGCTTCGAGCGTGGGGTGTGCGACGCGCAACCTCACCTTCTGTGTACATGCAG 3213  
QY 1039 AspProThrIleValArgIleGluProGluTrpSerIleValSerGlyAsnThrProIle 1058  
Db 3214 AACCGGTTCATCAGCGCATCAGTCCCGCCGCGCAGCTGTGTCAGTGGCGGAGGACATC 3273  
QY 1059 AlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGly 1078  
Db 3274 ACAGTGGCTGTGACCGTTTCCACATGTGTGAGATGTGTCCATGCGCGCTCCACCACT 3333  
QY 1079 GlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAla 1098  
Db 3334 GGCGGGGAGCC---ACGCTCTGCAAGGTTCCTCACTCCACCTCATCATCCTGCGCGTCC 3390  
QY 1099 ProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPhe 1118  
Db 3391 CCCGGGGCCCTGAGCAACGCATCAGCGCAGTGGACTTCTTCATCANTGGCGGGGCTTAC 3450  
QY 1119 GlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeu-----Asn 1132  
Db 3451 GCAGCAGAGGTGGCTGTGCTGAGGAGCTACTGGACCCCGAGGAGGCACACGGCGGAGC 3510  
QY 1133 LysThrAsnPheThrTyrTyrProAsnProValPheGluAlaPheGlyProSerGlyIle 1152  
Db 3511 AGTTCCTCGCTGGACTACCTCCCAACCACTGCTTCTACCGCCCAAGAGGGAGAGTGG 3570  
QY 1153 LeuGluLeuLysProGlyThrProIleIleLeu-----LysGlyLysAsnLeuIle 1169  
Db 3571 ATCAAGCACCACCCCGGAGGCTCTCACCTCGTTATCCACAGGAGCAGGACGCTG 3630  
QY 1170 ProProValAlaGlyGlyAsnValLysLeuAsnTyrThrValLeuValGlyGluLysPro 1189  
Db 3631 -----GGGCTCCAGAGTCAAGTACCGGTGCAAGTAGGCAAGTGAAGC 3675  
QY 1190 CysThrValThrValSerAspValGlnLeuLeuCysGluSerProAsn----- 1205  
Db 3676 TCGACATCCAGATCTCTCTGACAGANTCATCCACTGCTCGGTCAACGAGTCCCTGGGC 3735  
QY 1206 ---LeuIleGlyArgHisLysValMetAlaArgValGlyGlyMetGluTyrSerProGly 1224  
Db 3736 GCGCGCGTGGGAGCTGCCATCACATCCAGGTAGGGAACCTCAACACGACATCGCC 3795  
QY 1225 MetValTyrIleAlaProAspSerProLeuSerLeuProAlaIleValSerIleAlaVal 1244  
Db 3796 ACACGTGACAGTGGGGGAGC-----GAGACGGCCCATCATGCTGTCTCATCGTCATC 3846  
QY 1245 AlaGlyGlyLeuLeuIleIlePheIleValAlaValLeuIleAlaTyrLysArgLysSer 1264  
Db 3847 TCGAGGCTCCTGCTGCTGCTCTCCGTGGTGCC---CTGTTCGTCTTCTGTACCAAGAGC 3903  
QY 1265 ArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArg 1284  
Db 3904 CGACGTGCTGAGGTTACTGGCAGACAGCCTGCTGCAGATGGAGAGATGGAAATCTCAG 3963  
QY 1285 ValAlaLeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIleHisGluLeuThr 1304  
Db 3964 ATCCGAGAGAAATCCGAAAGGCTTCGCTGAGCTGCAGACAGACATGACAGATCTGACC 4023  
QY 1305 SerAspLeuAsp---GlyAlaGlyIleProPheLeuAspTyrArgThrTyrThrMetArg 1323  
Db 4024 AAGGAGCTGAACCGCAGCCAGGATCCCTCTCTGTGAGTATAAGCACTTCTGTCACCGC 4083  
QY 1324 ValLeuPhePro----- 1327  
Db 4084 ACCTTCTCCCAAGTGTCTCCCTTTATGAGAGAGCTTACGTGCTGCCCTCCACAGACC 4143  
QY 1328 -----GlyIleGluAspHisProValLeuArgAspLeuGlu 1339  
Db 4144 CTCACTCCCGGAGCAGCTCCAGGCACAGGAACCCACCCACTGCTGGGAGATGGAGG 4203



DR P-PSDB; ABG21017.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX  
 PS Claim 1; SEQ ID No 21008; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 6561 BP; 1329 A; 2077 C; 2004 G; 1130 T; 21 other;

Alignment Scores:  
 Pred. No.: 3,97e-184 Length: 6561  
 Score: 2357.50 Matches: 644  
 Percent Similarity: 49.85% Conservative: 345  
 Best Local Similarity: 32.46% Mismatches: 725  
 Query Match: 23.60% Indels: 271  
 DB: 23 Gaps: 62

US-09-964-956-13 (1-1896) x AAS85204 (1-6561)

QY 24 ThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPhe 43  
 DB 119 ACCCTGCTGGCGCTGCTGGCGAGGTGCCAGCTGAGCCCGCAGCTGGACTTCTC 178  
 QY 44 ArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIle 63  
 DB 179 CGCAGCCAG-----AAAGAGCTGACACACCTGCTGTGGATGAGGCTCAGGCGTGTG 232  
 QY 64 TyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThr 83  
 DB 233 TACCTGGGGCGGTGAATGCCCTACAGCTGATGCGAAGCTGCAGCTGGAGCAGCAG 292  
 QY 84 HisGluThrGlyProAspGluAspAsnProLysCysTyrProArgIleValGln-Th 103  
 DB 293 GTGCCACCGGGCGCCCTGGGACAAAGAAGTGCACGCCGCCCATCGAGGCCAGCCCA 352  
 QY 103 rCysAsnGluProLeuThrThrAsnAsnValAsnLysMetLeuLeuIleAspTyrLy 123  
 DB 353 GTGCCATGAG---CCTGAGATGACTGACATGTCAACCCACTGCTGTGTCGACCTCC 409  
 QY 123 sGluAsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIle-CysLysLeuLeuArgL 143  
 DB 410 CAGGAACCGCTGTGTGAGTGCAGGCGAGCTTCTTAAGGCGATCCTGCGCTCTGCGCGCC 469  
 QY 143 euGluAsp-----LeuPheLysLeuGlyGluProTyrHisLysLysLysLeuLeu 161  
 DB 470 TGACACACATCTCCCTCCGCTGTTCTACGAGGAGCGGCGGGAAGTCTTTCGTGG 529  
 QY 161 erGlyValAsnGluSerGlySerValPheGlyValIleValSerTyrSerAsnLeuAspA 181  
 DB 1517 CCTGCACCCAGTGCCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGAGGAC 1576

DB 530 CCAGCAATGATGAGGGCGTGGCCACAGCTGGGCTGGTGGCTCCAGGGTCTCTGGTGGT 589  
 QY 181 sPLysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerS 201  
 DB 590 ACCGGTCTGTTTGGGCAAGCAATGGCCACACAGCAACGGCATCATCGTGAGCA 649  
 QY 201 erArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAsp 221  
 DB 650 CTGGCTCTGTGACCGGACTACAGCAGGAGGCGCTTGAAGCGTACACAGGACACGCCA 709  
 QY 221 luPheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheA 241  
 DB 710 CCTCAAGCGCGGTACCTGCCACACACACAGCAGTTC----- 750  
 QY 241 spIleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnP 261  
 DB 751 -----GTGGCGGCTTTCGAGGAGGCGCCCTACGCTCTTGTGCTTCAACACG 799  
 QY 261 roGluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuV 281  
 DB 800 AGGACAAG---CACCGCGCCCGGAACCGCACG-----CTGCTGG 835  
 QY 281 alArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluValProIleGlyCys 301  
 DB 836 CAGCATGTGCAGAGAAGACCCACTACTCTCTACCTGGAGATGGAGCTGCAGTGCC 895  
 QY 301 luArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAla 321  
 DB 896 GGGACCCCGACATCCAC-----GCCCTGCCTTT-----GGCACCT 931  
 QY 321 alLeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerL 341  
 DB 932 GCCTGGCGGCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 991  
 QY 341 ysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeu 361  
 DB 992 GAGACAGC-----CGGACAGTGGGGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1045  
 QY 361 ysGlnIleAsnAspArgLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThrL 381  
 DB 1046 ACAAGGTGCAGCCCAAGATGGAGGCCAACCGAACCGCTGTGTAC-----ACAGG 1099  
 QY 381 euAspLeuAlaThrLeuLysValLys-----AspIleProCysSerSerAlaL 397  
 DB 1100 GGGAGCGCGGTGACATCTTCTACAGCCCTTCACGGCGATATCCAGTGGCGGCGCCAC 1159  
 QY 397 euLeuThrIleAspAspAsnPhe---CysGlyLeuAsp---MetAsnAlaProLeuGly 415  
 DB 1160 CGCCGGGTCCAGCAAGAGCTTCCCATGTGGCTCGGAGCAGCTGCCTACCCCTGGGCA 1219  
 QY 415 alSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThrSer 435  
 DB 1220 CGCGCGCGGGTTCAGAGCGCAGCGCTGCTGCGAGCGCTGAGAGCGCTGAGAGCGG 1279  
 QY 435 alIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLys 455  
 DB 1280 TGACGTGCGCGCGGAGAC 1339  
 QY 455 euLysLysIleArgValAspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGln 475  
 DB 1340 TCCTCAAGGTGATCCTCAC---CCAGATGGCAGCTCCCTCAGAGTAGTACGACTATC 1396  
 QY 475 alValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeu 495  
 DB 1397 TGAGATAAACAGAGAGTCAAGCGCGACCTGTGTGTGAGAGCTGGGAGCGCTGT 1456  
 QY 495 TyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrGln 515  
 DB 1457 AGCCATGACCCAGGACAGGTGTTCCGCTGCGGTGCGAGGTGCCTCAGTACCCGA 1516  
 QY 515 erCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTyrCysValLeuHisAsn 535  
 DB 1517 CCTGCACCCAGTGCCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGAGGAC 1576

QY 535 hrCysThrArgLysGluArgCysGluArgSerLysGluProArgArgPheAla---SerG 554  
|||||  
Db 1577 GATGACCGGAGCGCGAGTGTCCGCGCGGAGGAGGCCACCCACTGCTGGAGCC 1636  
QY 554 luMetLysGlnCysValArgLeuThr---ValHisProAsnAsnIleSer---ValSerG 572  
|||||  
Db 1637 GAAGCAAGTCTCTGCGCGCGTACACAGGCGCCAGCAGCAGACAGATGAGCGCGGCGCC 1696  
QY 572 lnTyraAsnValLeuLeuValLeuGluThrTyraAsnValProGluLeuSerAla-----G 590  
|||  
Db 1697 AGGGGAGGTGAGCTGACCGTACGCCCC-----CTCCCTGCCCCTGAGCGAGGAGAGC 1750  
QY 590 lyValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValGlnIle 610  
|||  
Db 1751 AGTTGCTGTCTTTTGGGAGTGCAGCGCCAGCAGCCCGCGCGTGGAGGCGGAGCGC 1810  
QY 610 leGlnCysTyraSerProAlaAlaLysGluValProArgIleLeuThrGluAsnGly---A 629  
|||  
Db 1811 TCATCTGCAACTCCCAACAGCAGC-----ATCCCC-----GTACACCCCGCAGGCGCAG 1858  
QY 629 sPHisHisValValGlnLeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerT 649  
|||||  
Db 1859 ACCAGTGGCGGTGACCATCCAGCTCTCTTACAGCAGGACACATCTTCTCAGCTCCT 1918  
QY 649 hrSerPheValPheTyraAsnCysSerValHisAsnSer-----CysL 663  
|||  
Db 1919 ACCAGTACCCCTTCTACGACTCGCGCCAGCCATGAGCTGGAGGAGAACCTGCCGTGCA 1978  
QY 663 euSerCysValGluSerProTyraArgCysHisTrpCysLysTyraArgHisValCysThrH 683  
|||||  
Db 1979 TCTCTGCGTGAACACCGCTGAGCTGCGGAGGACCTGCGCTACACGAGTGCAGG 2038  
QY 683 is---AspProLysThrCysSerPheGlnGluGly-----ArgValLysLeuProGluA 700  
|||  
Db 2039 AGGCTTCGCCCAACCT-----GAGACGGCATGCTCCGTGCCCATGAGGACA 2089  
QY 700 sCysProGlnLeuLeuArgValAspLysIleLeuValProValGluValIleLysProI 720  
|||||  
Db 2090 GCTGTCCCGACTTCTCTGGGACCGCAGCCCTGCTGATCCCATGAACACGAGCAGATG 2149  
QY 720 leThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGlnArgGlyTyraGlyCysI 740  
|||  
Db 2150 TGAATTCAGGCGCAAGAACCTGGAC-----2175  
QY 740 leLeuAsnIleGlnGlySerGluGlnArgValProAla-----LeuArgPheAsnSers 758  
|||||  
Db 2176 -----ACCGTGAAGGTCTCTCCCTGCAGTGGGAGTGCCTTCAAGTTCATGGAGC 2230  
QY 758 erSerValGlnCysGlnAsnThrSerTyraGluGlyMetGluIleAsnAsn-----776  
|||  
Db 2231 CGGTGACCATGAGGAATCTGGGACCTTGGCCCTTTCGGACCCCAAGCTCTCCACCATG 2290  
QY 777 -----LeuProValGluLeuThrValValTrpAsnGlyHisPheAsnIleAspA 793  
|||||  
Db 2291 CCAACGAGAGCTGCGCCCTGACCTCTAGCTCAAGTTCATGCGCAAG---AATATCGAC- 2346  
QY 793 snProAlaGlnAsnLysValHis-----LeuTyraLysCysGlyAlaMetArgGluSerC 811  
|||  
Db 2347 -----ACCAAGCTCCATGTGACCTCTACAACTGCTCTTGGCGCGGAGCGACT 2395  
QY 811 ysGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTyraCysGlnGlyProGlyG 831  
|||||  
Db 2396 GCAGCCTGTGCGGCGCGCTAACCAGCTACAGGTGTGCTGCGGCGGCGCAGACA 2455  
QY 831 lnCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuGluLeuSerGlyA 851  
|||  
Db 2456 GGTGCTGTATGAGGCCCTGTGCT-----AACA 2482  
QY 851 lalysSerLysCysThrAsnProArgIleThrGluIleProValThrGlyProArgG 871  
|||||  
Db 2483 CCACCTCCGAGTCCCGCGCGCTCATCACCAGGATCCAGCTGAGACGCGGCGCCCTG 2542

QY 871 luGlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPheArgAspIleA 891  
|||||  
Db 2543 GTGGGGCATCCGCATCACCATCCATCTGGGTCCAAATTTGGCGTCCCAAGCAGGAGCATC- 2601  
QY 891 laSerHisValLysValAlaGlyValGluCysSerProLeuValAspGlyTyraIleProA 911  
|||  
Db 2602 --CAGAGATCTCTGTGGCGCGCGGAACCTGCTCTTTTCAGCCGGAACCTTACTCGGTG 2659  
QY 911 laGluGlnIleValCysGluMetGlyGluAlaLysProSerGlnHisAlaGlyPheValG 931  
|||  
Db 2660 CCACCGGATCTGTGTGTATC---GAGGCTCGGAGACGCTTTCACGCGGGGTGTGTC 2716  
QY 931 luIleCysValAlaValCysArgProGluPhe-----MetAlaArgSerSer-----946  
|||  
Db 2717 AGTGGAGCTC-----TTCGGGAAACTGGGCGCTTCGCCTCCCAATG 2758  
QY 947 --GlnLeuTyraPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProM 966  
|||  
Db 2759 TCCAGTTACCTTCCACAGCCCAAGCCTCTCAGT---GTGGAGCCGACAGGAGCCGC 2815  
QY 966 etSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySerAsnVal 986  
|||  
Db 2816 AGCGGGGGGACACACTGACCATCCAGCGCACCCACTGACAGGCTCCAGGAG 2875  
QY 986 alValMetPheGlyLysGlnProCysLeuPheHisArgSerProSerTyraIleValC 1006  
|||  
Db 2876 ACGTG-----CGGTGACCTCAACGCGTCCCGT 2905  
QY 1006 ysAsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnValAspArgA 1026  
|||  
Db 2906 GTAAAGTGACGAAGTTTGGGCGCAGCTCCAGTGTGTCACTGGCCCCCAGGACAGCGG 2965  
QY 1026 laLysIle-----HisGlnAspLeuValP 1034  
|||  
Db 2966 GCCAGATGCTTCGGAGTCTCTACGGGGGTCCCGCTGCCCAACCCCGCATCTTCT 3025  
QY 1034 heGlnTyraValGluAspProThrIleValArgIleGluProGluTrpSerIleValSerG 1054  
|||  
Db 3026 TCACCTACCGGAAACCCCTACTGCGAGCCTTCGAGCGCTACGAGCTTTGCCAGT 3085  
QY 1054 lyAsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleA 1074  
|||  
Db 3086 GTGCGCGCAGCATCAACCTCAGCGGTTCAGGCTTCAGGCTGTATCCAGAGTTTGGCATG- 3144  
QY 1074 rgAlaLysHisGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluM 1094  
|||  
Db 3145 -----GTGGTATCGGAGGCCCTCGAGTCTCTGCGCAGC 3178  
QY 1094 etThrCysGlnAla-----ProAlaLeuAlaLeuGlyProAspHis-----1107  
|||  
Db 3179 CGCGCGGAGGCTCAATCCCTGCGAGCCCATGACGCTGTGGGTACAGACTACGTGTTC 3238  
QY 1108 -----GlnSerAspLeuThrGluArgProGluGluPheG 1119  
|||  
Db 3239 ACAATGACACCAAGTCTCTTCTGCTCCCGGTGTGCTGAGGAGCCAGAGGCTTACAG 3298  
QY 1119 lyPheIle-----LeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrA 1135  
|||  
Db 3299 ACCTCAGGCTGATCCAGATGAGCGGACCGCTGCTGCTCAGACAGAGCGCGGG 3358  
QY 1135 snPheThrTyraTyraProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluL 1155  
|||  
Db 3359 CTTTCGAGTACGTGCTGACCCACTTTGAGAACTTC-----3396  
QY 1155 euLysProGlyThrProIleLeuLysGlyLysAsnLeuIleProProValAlaGlyG 1175  
|||  
Db 3397 -----ACAGGTGGGTCAAGAAGAGGTCAACAAGCTCATCCGCGCCCGGGCA 3445  
QY 1175 lyAsnValLysLeuAsnTyraThrVal-----LeuValGlyGluLysProC 1190  
|||  
Db 3446 CCANTCTGAACAGCGGATGACGTGTCAGGAGCGCCCTTCGTGGGTGCGGAGCGCT 3505  
QY 1190 ysThrVal---ThrValSerAspValGlnLeuLeuCysGluSerProAsnLeuIleGly- 1208

Db 3506 GCACCATGAAGACGCTGACGGAGAACGACCTGTACTGTGAGCCGCCCGGAAGTTGCAGCCC 3565  
Qy 1209 -----ArgHisLysValMetAlaArgValGlyGlyMet-Glu 1220  
Db 3566 CCGCCCAAGCGGGGAGAACAGAGACACACACAAACCTGCCCGAGTTCAATTGTGAA 3625  
Qy 1221 TyrSerProGlyMetValTyrIleAlaPro----- 1230  
Db 3626 GTTCGGCTCTCGGAGTGGGTGGCGCGGTGGAGTACGACACACGGGTGAAGCGACG 3685  
Qy 1231 ---AspSerProLeuSerLeuProAlaIleValSerIleAlaValAlaGlyGlyLeuLeu 1249  
Db 3686 TGCGGCTCAAGCCTCTCTTCGCGCTGGTCATC-----GTGCC 3724  
Qy 1250 IleLeuPheIleValAlaValLeuIle---AlaTyrLysArgLysSerArgGluSerAsp 1268  
Db 3725 ATGGTGTCTGTCATCGCGGTGCTGTCTACTGTCTGGAGAGAGCCAGCAGCCGAA 3784  
Qy 1269 LeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuGluSerArgValAlaLeuGlu 1288  
Db 3785 CGAGAGTATGAGAAGATCAAGTCCACGCTGAGGGCCTGGAGGAGAGCGTGGGACCGC 3844  
Qy 1289 CysLysGluAlaPheAlaGluLeuGlnThrAspIleHisGluLeuThrSerAspLeuAsp 1308  
Db 3845 TGCAAGAAGGAATTCACAGACCTGATCATCGAGATGGAGGACCACCAACAGCAGCTGCAC 3904  
Qy 1309 GlyAlaGlyIleProPheLeuAspTyrArgThrTyrThrMetArgValLeuPhe----- 1326  
Db 3905 GAGCGCGGATCCCGCTGCTGGACTACAGACCTACACCGCCGCTCTTCTCTCTGCC 3964  
Qy 1327 -----ProGlyIleGluAspHisProValLeuArgAspLeuGluValProGlyTyrArg 1344  
Db 3965 TCCAAAGGACGCGACAAGGACGTATGATACCGCGCAAGCTGGACATCCCTCGCGCGG 4024  
Qy 1345 GlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIleAsnAsnLysValPhe 1364  
Db 4025 CGCGCGGTGGAGAGCGCCTTACACAGTCTCCAACTCTGGAACAGCAAGTCTTTC 4084  
Qy 1365 LeuLeuSerPheIleArgThrLeu-GluSerGlnArgSerPheSerMetArgAspArgG1 1384  
Db 4085 CTATCATATTCATCCACACCTCGGAGAACCCGCGGAGTCTCGCGCCGCGCAAGGT 4144  
Qy 1384 yAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeuGluTyrAlaThrAspVa 1404  
Db 4145 CTACTTCGCTCCCTGCTGAGCGTGGCGTGCACGGAACTGGAGTACTACAGGACAT 4204  
Qy 1404 lLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGluSerLysAsnHisProLy 1424  
Db 4205 CATGCACACGCTTCTCTGGAGCTCTGGAGCAGTACGCTGCGCAAGAAC---CCCAA 4261  
Qy 1424 sLeuLeuLeuArgArgThrGlu-SerValAlaGluLysMetLeuThrAsnTrpPheThrP 1444  
Db 4262 GCTGATCTGCGCAGGTCTGTATCTGTGTGGAGAGGATGCTGTCCAACTGGATGTCCA 4321  
Qy 1444 heLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeuPheSerLeuPheCysA 1464  
Db 4322 TCTGCTGTACCATCTAAGGACAGTGTGCGGGAGCCCTGTACAAAGCTCTTCAAG 4381  
Qy 1464 lalLeuGlnGlnMetGluLysGlyProIleAspAlaIleThrGlyGluAlaArgTyrS 1484  
Db 4382 CCATCAACATCAGTGTGAAGAGCGCGGTGGATGCGGTACAGAGAGGCCAAGTACA 4441  
Qy 1484 erLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLysThrLeuValLeuSerC 1504  
Db 4442 CTCTCAACGACACGGGCTGTGGGGATGATGTGGAGTACGACGCCCTCGCTGAGCG 4501  
Qy 1504 ysValSerProAspAsnAlaAsnSerProGluValProValLysIleLeuAsnCysAspT 1524  
Db 4502 TGATCGTGCAGGACGAG---GGAGTGGACGCCATCCCGGTGAAGGTCTCAACTGTGACA 4558  
Qy 1524 hrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLysAsnValProCysSerH 1544

Db 4559 CCATCTCCCGAGGTCAAGGAGAGATCATTTGACCAGGTGTACCGTGGGCGACCCCTGCTCT 4618  
Qy 1544 isArgProLysAlaAlaAspMetAspLeuGluTrpArgGlnGlySerGlyAlaArgMetI 1564  
Db 4619 GCTGCCCGAGCCAGACACGCGTGGTCTCTGGAGTGGCGTCCGGCTCCACAGCGCAG---A 4675  
Qy 1564 leLeuGlnAspGluAspIleThrLysIleGluAsnAspTrpLysArgLeuAsnThrL 1584  
Db 4676 TCCTGTCCGACTGTGACCTGCAGCTCACAGCGGAGCGCGTGGAGAGCGCTCAACACC 4735  
Qy 1584 euAlaHisTyrClnValProAspGlySerValValAlaLeuValSerLysGlnValThrA 1604  
Db 4736 TTATGATCATCAATGTCCGGATGAGCACCCCTCATCTCTG----- 4776  
Qy 1604 laTyrAsnAlaValAsnAsnSerThrValSerArgThrSerAlaSerLysTyrGluAsnM 1624  
Db 4777 -----TCCAAGGTGGGCTCTCCAG----- 4797  
Qy 1624 etIleArgTyrThrGlySerProAspSerLeuArgSerArgThrProMet-----IleT 1642  
Db 4798 -----CAGCGGAGGACAGCGCAAGAACTGCCCTGGGAGCGCCA 4840  
Qy 1642 hrProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHisGluHisGlyA 1662  
Db 4841 TGCCCTC-CTGAGGAGGAGAACCGGTGTGTGCACCTGTGTGGCCGACCGACGAGTGG 4899  
Qy 1662 spGlnLysGluGlyAspArgGlySer-----LysMetValSerG 1675  
Db 4900 ACGAGGGCAAGTCCAAGAGGACGCGTGAAGAGAGGAGCGGACGAAGCCATCACCG 4959  
Qy 1675 loIleTyrLeuThrArgLeuLeuAlaThr-LysGlyThrLeuGlnLysPheValAspAsp 1694  
Db 4960 AGATCTACTGTACGGCGCTCTCTCAAGTCAAGGCGACACTGCAGCAGTTGTGGACAAC 5019  
Qy 1695 LeuPheGluThrIlePheSerThrAlaHisArgLysSerAlaLeuProLeuAlaIleLys 1714  
Db 5020 TTCCTCCAGAGCGTCTGCGCCTGGGCAC-----GCGTGCACCTGTCAGTCAAG 5070  
Qy 1715 TyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisVal 1734  
Db 5071 TACTTCTTCGACTTCTTGACGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5130  
Qy 1735 ArgHisThrTrpLysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsn 1754  
Db 5131 ATCCACATCTGGAAGACGAAACAGCTTACCGCTCCGCTTCTGGGTGAACATCTCAAGAAC 5190  
Qy 1755 ProGlnPheValPheAspIleHis-LysAsnSerIleThrAspAlaCysLeuSerValVa 1774  
Db 5191 CCCCCTTCTTTCGCTGATGTCGACGAGGTGTCGACGAGGTGTCGCGCTGTCAGTCAT 5250  
Qy 1774 lAlaGlnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerPr 1794  
Db 5251 CGCGCAGACCTTCATGGATGCTGACCGGCGCAGGAGCATAGCTGAGCGCGGATCTCC 5310  
Qy 1794 oSerAsnLysLeuLeuTyrAlaLysAspIleProSerTyrLys-AsnTrpValClnuArgT 1814  
Db 5311 CAGCAACAGCTGCTGTACGCCAAGGAGATCTCCACTACAGAAGATGTTGGGAGGATT 5370  
Qy 1814 yr-TyrSerAspIle-GlyLysMetProAlaIleSerAspGln-Asp-MetAsnAlaTyr 1832  
Db 5371 ACTTACAAGGGATCCGGGCGATGTCGAGGTGTCGACGAGGAGGAGGAGGAGGAGGAGGAG 5430  
Qy 1833 LeuAlaGluGln---SerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeuSer 1851  
Db 5431 CTTCGACAGAGATTTTCCCGGGCGGACAGGACTCTTGAACACACCTCTGCGGACTCCAC 5490  
Qy 1852 GluIlePheSerTyrValGlyLysTyrSerGluGluIleLeuGlyProLeuAspHisAsp 1871  
Db 5491 CAGCTCTACCAATACACGAGAGTACTATGACGAGATCATCAATGCCTTGGAGGAGAT 5550  
Qy 1872 AspGlnCysGlyGlnLys-LeuAlaTyrLysLeuGluGlnVal 1886  
Db 5551 CCGTCCGCCCGAGAGATGAGCTTGGCTTCCGCTTCCGCTGACGAGATT 5596

Query Match:	18.55%	Indels:	597
DB:	22	Gaps:	72
US-09-964-956-13 (1-1896) x AAS000021 (1-5367)			
QY 25	LeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPheArg	44	
DB 43	CTGCTGTCGCCACCGCCACTGCTCCCTTGACAGGGGCCCATCGCTTC		
QY 45	GlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIleTyr	64	
DB 88	TCCGCACCTAATACCACTCTCAACCACTTGCACCTGGCACCTGGCGAGGCACACTCTAT	147	
QY 65	LeuGlyAlaValAsnArgIleTyrLysSerSerAspLeuLysValLeuValThrHis	84	
DB 148	GTCCGGCAGCTGAACCGCCCTTCCAGCTCAGCCCGCAGCTGCAGCTCGAGGCCGTGGCT	207	
QY 85	GluThrGlyProAspGluAspAsnProLysCysTyrProProArgIleValGlnThrCys	104	
DB 208	GTCACTGGCCCTGTAATCGACAGCCCTGACTGCGTGCCCTCCCGTACCCAGCCGAGTGC	267	
QY 105	AsnGluProLeuThrThrThrAsnAsnValAsnLysMetLeuLeuIleAspTyrLysGlu	124	
DB 268	CCACAG--GCCCAGCTCACTGACAAATCCCAACCACTGCTGCTGTTGAGCAGCCGCGCC	324	
QY 125	AsnArgLeuIleAlaCysGlySerLeuTyrGlnGlyIleCysLysLeuLeuArgLeuGlu	144	
DB 325	CAGGAGCTGTGGCTCGGGCAGGTGGCGGAGGCGGTGTGTGAGACACGGCGCCTTGGG	384	
QY 145	Asp-----LeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeu	160	
DB 385	GATGTGCCCGAGTGCTGTACCAGCTCAGACCCCT-----GGTGACGGCGAGTTT	435	
QY 161	SerGlyValAlaAsnGluSerGly---SerValPheGlyValIleValSerTyrSerAsnLeu	179	
DB 436	GTGGCTGCCAATACCCCGGAGTGGCAACGGTGGGGCTGGTGGTGCCTTGCCCGCC	492	
QY 180	AspAspLysLeuPheIleAlaThrAlaValAspGlyLys---ProGluTyrPheProThr	198	
DB 493	CGGAGCTCTCTGTGTGGCCAGAGGCTTGGCGGCAAGCTGTGCGCAGGGGTGCCACCC	552	
QY 199	IleSerSerArgLysLeuThrLysAsnSer-----GluAlaAspGlyMetPheAlaTyr	216	
DB 553	CTGGCCATCCGCACTGGCGGGTCTCAGCCCTTCTCCAGCAGGGCTGGCGCCCTTCG	612	
QY 217	ValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThrPheThrIle	236	
DB 613	GTGTGGCGAC-----CGCGGGCGCCCGCCAGCTCAG---TAC 723	624	
QY 237	IleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPhe	256	
DB 625	TTCTCCGACTACAAACAGCTAGCTCGGGCGCTTTGGCGAGCCCGCTCCGCCCTACTTC	684	
QY 257	LeuThrLeuGlnProGluMetValSerProProGlySerThrThrLysGluGlnValTyr	276	
DB 685	GTGTTCCCGCC-----CGCGGGCGCCCGCCAGCTCAG---TAC 723	686	
QY 277	ThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyrValGluVal	296	
DB 724	CGCTCTTACGTGGCCCGCGCTGCTGCTGGGGACACCAACCTGTACTCTTACGTGGAGGTC	783	
QY 297	ProIleGlyGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSer	316	
DB 784	CCCTTCCCTGCGCAGGCGCAGGCG-----CTCATCCAGCCCGCCTTCCTTGCC	831	
QY 317	LysAlaGlyAlaValLeuLeuGlyArgThrLeuGlyValHisProAspAspLeuPhe	336	
DB 832	---CCGGGCGCCTTGCTAGG-----849	849	
QY 337	ThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCys	356	
DB 850	---GTGTTTGGCGGGCGCCAAAGGCG-----ACCCAGCGCGCTCTGT 891	891	



Db 2568 GCGACTGATCGTGTGTCAGGGGACCCGCGCTAGACGTGGTGCGACGCGCCCTACTGTCTGT 2627  
Qy 1013 uValLeuGluMetLysValSerValGlnValAlaAspArgAlaLysIleHisGlnAspLeuVa 1033  
Db 2628 GTGGCTGGAGGTGACGACAGAGGTGACGCTTCCAGGGCCAG----- 2670  
Qy 1033 lPheGlnTyrValGluAspProThrIleValArg----- 1leGluPr 1047  
Db 2671 -----CCCCAGGACCCACAGCAAGAGGAGCTGTGTGAGCCCTCTCTCGGACCC 2720  
Qy 1047 oGluTrpSerIleValSerGlyAsn-----ThrProIl 1058  
Db 2721 CCAGGCTTGTATCCAGCTCGTGGGGGCTGTGACGGCCACAGACAGCCCGAGCTCACT 2780  
Qy 1058 eaLaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisG1 1078  
Db 2781 CCACCTGTGTGGCC-----CTGAATGCCCCACACAGTGTCC----- 2817  
Qy 1078 yGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAl 1098  
Db 2818 -----ACCGTGTCTCCGTCACCTGCTCCAGCCTCTCTGTGCGCGAG 2861  
Qy 1098 aProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPh 1118  
Db 2862 CCTGCTGTA-----CCAGACAGA-----GCCACCCGCGAGCGGT 2897  
Qy 1118 eGlyPheIleLeuAspAsnValGln-----SerLeuLeuIleLeuAsnLysThrAsnPh 1136  
Db 2898 CTTCTTCAACCTTAGACAACGTGCAAGTGAGCTTCGCGAGTGCCAGTGCGGGGCCAGGCGTT 2957  
Qy 1136 eThrTyrTyrProAsnProValPheGluAlaPhe-----GlyProSerGlyIleLe 1153  
Db 2958 CTTGTACACGACCAACCCCGCTGCGACCCCTCAGCCGCGAGGGGCTGCGCCGCCCTA 3017  
Qy 1153 uGluLeuLysProGlyThrProIleLeuLysGlyLysAsnLeuIleProValAl 1173  
Db 3018 CGCCTCAGCAGCGGCATGTCTGTGATGTGGAGGCGGAGGCTCAACCTGGGCATFCAG 3077  
Qy 1173 aGlyGlyAsnValLysLeuAsnTyrThrValLeuValGlyGluLysProCysthrVal-- 1192  
Db 3078 CAAGGAGGAGTGGCGGTGCAC-----ATCGCGCGCGGAGTGCTGTGTGA 3125  
Qy 1193 -ThrValSerAspValGlnLeuLysGlyLysPro-----As 1205  
Db 3126 GACGTTCAGCGCACACCTGTACTGTGCGAGCGCGCTGCGCACGCCCCGAGCGCTGCCAA 3185  
Qy 1205 nLeuIleGlyArgHisLysValMetAlaArgValGlyGlyMetGluTyrSerProGlyMe 1225  
Db 3186 TGGCTCCGCGTCCACAGTTGTTGTGTGAGATGGCAATGTGCGCTGGCCCTGGGCC 3245  
Qy 1225 tValTyrIleAlaProAspSerProLeuSer-----LeuProAlaIleValSerIl 1242  
Db 3246 TGTGAGTACGAGGCTGAACCCCGCTGTCTGCTTTCCCGTGGAGGCCCGAGCGCGT 3305  
Qy 1242 eaLaValAlaGlyGlyLeuLeuIlePheIleValAlaValLeuIleAlaTyrLysAr 1262  
Db 3306 GGGCATGGGTGTGCTGAGTGTGATCCCGCGTGTCTCTCCTCCTCATGCTATACAGGCA 3365  
Qy 1262 glySerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeuG1 1282  
Db 3366 CAAGACAGCAGGCGCTCGGGACTACAGAGAGGTGTAGTGCAGCTGGAGAGCTTGA 3425  
Qy 1282 uSerArgValAlaLeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIleHisG1 1302  
Db 3426 GACCGCGTGGGAGACAGTGGCGCAAGGAGTTACAGAGACCTCATGACGGAGATGACCGA 3485  
Qy 1302 uLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspTyrArgThrMe 1322  
Db 3486 CTTACAGCAGCAGCTGGAGGCGAGCGGATCCCTTCTTGACTACCGCACTACGCGCGA 3545  
Qy 1322 tArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLeuValProG1 1342  
Db 3546 GCGCGCTTCTTCCCTGGCCATGGCGGTGCGCGCTGCGAGCCCAAGCCTGAGGGGCGCAGG 3605

Qy 1342 yTyr-----ArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIleAs 1360  
Db 3606 GGAGCAGCGGCACCTGTGCACTGTGCGCGCAGGCGCTCAGCAGCTCTCCAACCTGCTCAA 3665  
Qy 1360 nAsLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPheSerMe 1380  
Db 3666 CAGAACTCTCTCTCTCTCAGC-----GTGAGGCGCTGTGGCGGAGTGCCCA----- 3714  
Qy 1380 tArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeuGluTyr 1400  
Db 3714 ----- 3714  
Qy 1400 rAlaThrAspValLeuLysGlnLeuAlaAspLeuIleAspLysAsnLeuGluSerLys 1420  
Db 3715 -----GTGGCAAGAGGTGGCGTGG 3737  
Qy 1420 sAsnHis---ProLysLeuLeuLeu---ArgArgThrGluSerValAlaGluLysMetLe 1438  
Db 3738 GAACCTACTGGCTGAGACAAAGGTGGGGGAGGAGACAGACCATGGTGGAGAAATGCT 3797  
Qy 1438 uThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLe 1458  
Db 3798 CACCACTGGCTGTCCATCTGCTGTGAGGAGGTGGTGTGTAACCACT 3857  
Qy 1458 uPheSerLeuPheCysAlaIleLysGlnGlnMetGluLysGlyProIleAspAlaIlePh 1478  
Db 3858 GTACATGCTCTTCGGGCGCATCCAGTACAGGTGACAAAGCCCGCTGACCGCTGAC 3917  
Qy 1478 rGlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLys 1498  
Db 3918 AGGCAAGGCCAAACGACCCCTCAATGATAGCGCTTGTGCGGAGGAGCTGAGTTCCA 3977  
Qy 1498 sThrLeuValLeuSerCysVal-----SerProAspAsnAlaAs 1511  
Db 3978 GCCCTGACGCTGATGTGTGGCGCGCGGCTGGCGGGCGCGAGCAGCAGCGCA 4037  
Qy 1511 nSerProGluValProValLysIleLeuAsnCysAspThrIleThrGlnValLysGlu 1531  
Db 4038 GATCAGCGCGCTGCCAGCGCGGTCTCGACACGACACCATCACCAGGTCAAGGAGAA 4097  
Qy 1531 sIleLeuAspAlaIlePheLysAsnValProCysSerHisArgProLysAlaAlaAspMe 1551  
Db 4098 GGTGTGACCAAGTCTACAAGGCGACCCCTCTCCAGAGGCGCTCAGTGCATGCCCT 4157  
Qy 1551 tAspLeuGluTrpArgGlnGlySerGlyAlaArgMetIleLeuGlnAspGluAspIlePh 1571  
Db 4158 AGACCTGGTGAGAGCGCGCCCT----- 4182  
Qy 1571 rThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHisTyrGlnValProAs 1591  
Db 4183 -----GCCAC-----CCACCCCA 4196  
Qy 1591 pGlySerValValAlaLeuValSerLysGlnValThrAlaTyrAsnAlaValAsnAsnSe 1611  
Db 4197 GGGACCCCTTCCCTACCCCTCCCGCA----- 4221  
Qy 1611 rThrValSerArgThrSerAlaSerLysTyrGluAsnMetIleArgTyrThrGlySerPr 1631  
Db 4222 -----CCTGGAGCCCC 4232  
Qy 1631 oAsp---SerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyVal 1650  
Db 4233 TCAACTGTGTCTTACTATGAACATACCCAGCGTGGAGGATGCGAGGAGGGGGGTGTG 4292  
Qy 1650 sMetTrpHisLeuValLysAsnHisGluHisGlyAsp----- 1662  
Db 4293 CCTTGGCACCTGTGTAAGGCCACAGGAGGCCAGAGGCCAGAGGGGCCAAGGTGCGGTGACGAC 4352  
Qy 1663 -----GlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLeuThrAr 1680  
Db 4353 CCTCGGAGGCGCGAGCCAGCAAGG---GCCAAGCCATTCGGGAATCTACTCTCACCCG 4409

[illegible]

CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention.

XX Sequence 1088 BP; 256 A; 339 C; 310 G; 183 T; 0 other;

## Alignment Scores:

Pred. No.: 1,95e-120 Length: 1088  
Score: 1573.50 Matches: 300  
Percent Similarity: 90.70% Conservative: 22  
Best Local Similarity: 84.51% Mismatches: 32  
Query Match: 15.75% Indels: 1  
DB: 22 Gaps: 1

US-09-964-956-13 (1-1896) x ABA08997 (1-1088)

QY 1420 LysAsnHisProLysLeuLeuLeuArgArgThrGluSerValAlaGluLysMetLeuThr 1439  
DB 27 AAGACCACCCCAAGCTGCTACTGCCCGGACTGAGTGGCGCAGAGAAGATGCTAACT 86  
QY 1440 AsnTrpPheThrPheLeuLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeuPhe 1459  
DB 87 AACTGGTTCACCTTCCTCTGTATAGTTCTCAAGGAGAGCGCTGGGAGCGCTGTTT 146  
QY 1460 SerLeuPheCysAlaIleLysGlnMetGluLysGlyProIleAspAlaIleThrGly 1479  
DB 147 ATGCTGTACTGGCCATCAAGCACCAGATGGAGAGGGCCCCATTGACGCCATCAGGGT 206  
QY 1480 GluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLysThr 1499  
DB 207 GAGGACGCTACTCCTGAGTGAGGACAAGCTATTCTGGCACCCTGATTGACTACAGACA 266  
QY 1500 LeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLysIle 1519  
DB 267 CTGACCTCAACTGTGTGAACCTGAGATGAATGAATGACCTGAGTGGCGGTGAAGGG 326  
QY 1520 LeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLysAsn 1539  
DB 327 CTGGACTGTGACAGGGGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 386  
QY 1540 ValProCysSerHisArgProLysAlaAlaAspMetAspLeuLeuTrpArgGlnGlySer 1559  
DB 387 GTGCCCTACTCCAGCGGCGCCAGGCGCGGACATGACCTGAGTGGCGCCAGGCGCG 446  
QY 1560 GlyAlaArgMetIleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTrpLys 1579  
DB 447 ATGGCGCGCATCTCTCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 506  
QY 1580 ArgLeuAsnThrLeuAlaHisTyrGlnValProAspGlySerValValAlaLeuValSer 1599  
DB 507 AGGCTGAACACACTGGCTCACTACCAAGTGACAGCGGGTCTCTCGGTGGCAGCTGGTCC 566  
QY 1600 LysGlnValThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSerAlaSer 1619  
DB 567 AAGCAGACCTCGCGCTACACATCTCCACTCTCCACTTCCACCAAG---TCCCTCAGC 623  
QY 1620 LysTyrGluAsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgThrPro 1639  
DB 624 AGATACGAGAGCATGTGCGCAGCGGCGCAGCAGCGGCGGCGGCGGCGGCGGCGG 683  
QY 1640 MetIleThrProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHisGlu 1659  
DB 684 ATGATCACCCGACTGTGAGAGCGGCGCACCACCAAGCTGTGGCAGTGTGAGAACACGAC 743  
QY 1660 HisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLeuThr 1679  
DB 744 CACCTGGACACCGGTGAGGCTGACCGCGCAGCAAGATGGTCTCGGAGATCTACTTGACA 803  
QY 1680 ArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspAspLeuPheGluThrIle 1699  
DB 804 CGGCTACTGGCCACCAAGGCGACACTGCGAGAAGTTTGTGACGACCTGTTTGAGACCATC 863  
QY 1700 PheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPhe 1719

DB 864 TTCAGCACGGCACACCGGGGCTCAGCCCTGGCGTGGCATCAAGTACATGTTGACTTC 923  
QY 1720 LeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrpLys 1739  
DB 924 CTGGATGAGCAGGCGGACAGCAGATCCACGATGCTGAGCTGCGCCACACCTGGAAG 983  
QY 1740 SerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheValPhe 1759  
DB 984 AGCAACTGCCTGCCCTGCGCTTCTGGGTGAAGTGATCAAGAACCCACAGTTTGTGTTT 1043  
QY 1760 AspIleHisLysAsnSerIleThrAspAlaCysLeuSerValVal 1774  
DB 1044 GACATTCAAGAAGACAGCATCAGCAGCGCTGCTTGTGGTGGTG 1088

## RESULT 15

AAK94342  
ID AAK94342 standard; cDNA: 2971 BP.

XX AAK94342;

AC 07-NOV-2001 (first entry)

XX Human cDNA, SEQ ID NO: 3041.

DE Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EP1130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-0114089.

XX 08-JUL-1999; 99JP-0194486.

XX 11-JAN-2000; 2000JP-0118774.

XX 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

XX P-PSDB; AAM93421.

XX 830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation -

XX Disclosure; SEQ ID NO 3041; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a human cDNA provided in the specification.

XX Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 2971 BP; 647 A; 921 C; 894 G; 509 T; 0 other;

## Alignment Scores:

Pred. No.: 8.78e-114 Length: 2971  
Score: 1501.00 Matches: 320  
Percent Similarity: 59.72% Conservative: 150  
Best Local Similarity: 40.86% Mismatches: 225  
Query Match: 15.03% Indels: 92

[illegible]

```

Db      1950  TAGCCCAAGGAGATCTCCACCTACAGAGAGATCGTGGAGGATTACTACAGGGGATCCGG 2009
QY      1820  LysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluGlnSerArgMet 1839
      ::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      2010  CAGATGGTCAGGTACGCCAGCAGGACATGAACACACACACCTGGCAGAGATTTCCTCCGGCG 2069
      ::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      1840  HisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGlyLys 1859
      ::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      2070  CACACGGACTCCTTGAACACCCCTCGTGGCAGCTCCACAGCTCTACCAATACAGCAGAG 2129
      ::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      1860  TyrSerGluGluIleLeuGlyProLeuAspHisAspGlnCysGlyLysGlnLysLeu 1879
      ::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      2130  TACTATGACGAGATCATCAATGCTTGGAGGAGGATCCTGCCGCCAGAGATGCAGCTG 2189
      ::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      1880  AlaTyrLysLeuGluGlnVal 1886
      ::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      2190  GCCTTCCGCTGCAGCAGATT 2210
      ::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

Search completed: June 28, 2003, 11:41:46  
 Job time : 1613 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 07:26:01 ; Search time 173 seconds  
(without alignments)  
3361.035 Million cell updates/sec

Title: US-09-964-956-13  
Perfect score: 9990  
Sequence: 1 MKAMPNWTCLLSHLLVMG.....OKLAYKLEQVITLMSLDSNK 1896

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09964956/runat\_24062003.124340.5230/app\_query.fasta\_1.2055  
-DB=Issued Patents\_NA -OFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09964956.ecgn\_1.1.78 @runat\_24062003.124340.5230 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1349	13.5	4707	3	US-09-181-706-1
2	1349	13.5	4707	4	US-09-458-791-1
3	1349	13.5	4707	4	US-09-459-066-1
4	629.5	6.3	4626	1	US-08-306-691B-22
5	629.5	6.3	4626	5	PCT-US93-06251-27
6	227	2.3	2670	1	US-08-121-713D-61
7	227	2.3	2670	1	US-08-835-268-61
8	227	2.3	2670	2	US-09-060-692-61
9	227	2.3	2670	3	US-08-833-391-61
10	227	2.3	2670	4	US-09-060-610-61
11	227	2.3	2670	5	PCT-US94-10151A-61
12	207	2.1	2854	1	US-08-121-713D-57

13	207	2.1	2854	1	US-08-835-268-57	Sequence 57, Appl
14	207	2.1	2854	2	US-09-060-692-57	Sequence 57, Appl
15	207	2.1	2854	3	US-08-833-391-57	Sequence 57, Appl
16	207	2.1	2854	4	US-09-060-610-57	Sequence 57, Appl
17	207	2.1	2854	5	PCT-US94-10151A-57	Sequence 53, Appl
18	200	2.0	2601	1	US-08-121-713D-53	Sequence 53, Appl
19	200	2.0	2601	2	US-08-835-268-53	Sequence 53, Appl
20	200	2.0	2601	3	US-09-060-692-53	Sequence 53, Appl
21	200	2.0	2601	4	US-08-833-391-53	Sequence 53, Appl
22	200	2.0	2601	5	US-09-060-610-53	Sequence 53, Appl
23	200	2.0	2601	5	PCT-US94-10151A-53	Sequence 53, Appl
24	192.5	1.9	3560	1	US-08-121-713D-59	Sequence 59, Appl
25	192.5	1.9	3560	1	US-08-835-268-59	Sequence 59, Appl
26	192.5	1.9	3560	2	US-09-060-692-59	Sequence 59, Appl
27	192.5	1.9	3560	3	US-08-833-391-59	Sequence 59, Appl
28	192.5	1.9	3560	4	US-09-060-610-59	Sequence 59, Appl
29	192.5	1.9	3560	5	PCT-US94-10151A-59	Sequence 59, Appl
30	186.5	1.9	2504	1	US-08-121-713D-63	Sequence 63, Appl
31	186.5	1.9	2504	2	US-08-835-268-63	Sequence 63, Appl
32	186.5	1.9	2504	2	US-09-060-692-63	Sequence 63, Appl
33	186.5	1.9	2504	3	US-08-833-391-63	Sequence 63, Appl
34	186.5	1.9	2504	4	US-09-060-610-63	Sequence 63, Appl
35	186.5	1.9	2504	5	PCT-US94-10151A-63	Sequence 63, Appl
36	195	1.6	2898	4	US-09-308-179B-2	Sequence 2, Appl
37	149.5	1.5	2498	4	US-09-041-236-1	Sequence 1, Appl
38	148	1.5	2010	4	US-09-240-410-1	Sequence 1, Appl
39	146.5	1.5	1818	4	US-09-041-236-3	Sequence 3, Appl
40	131	1.3	4435	1	US-08-308-872B-5	Sequence 5, Appl
41	130	1.3	2050	4	US-09-026-001A-5	Sequence 5, Appl
42	130	1.3	2972	1	US-08-454-455-3	Sequence 3, Appl
43	129	1.3	2481	3	US-09-188-930-64	Sequence 64, Appl
44	128	1.3	4758	3	US-09-191-647-1	Sequence 1, Appl
45	128	1.3	4758	4	US-09-540-245A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-181-706-1  
; Sequence 1, Application US/09181706  
; Patent No. 6130088  
; GENERAL INFORMATION:  
; APPLICANT: Melanie K. Spriggs, Michael R. Comeau,  
; APPLICANT: Robert F. DuBoise, Richard S. Johnson  
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN  
; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Janis C. Henry  
; STREET: 51 University St.  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/181,706  
; FILING DATE: October 28, 1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/958,598 (converted to a  
; APPLICATION NUMBER: Provisional, see below)  
; FILING DATE: October 28, 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: --to be assigned-- (USSN 08/958,598  
; APPLICATION NUMBER: conversion to Provisional application)  
; FILING DATE: October 26, 1998  
; CLASSIFICATION:

## ATTORNEY/AGENT INFORMATION:

NAME: Henry, Janis C  
 REGISTRATION NUMBER: 34,347  
 REFERENCE/DOCKET NUMBER: 2631-A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206)470-4189  
 TELEFAX: (206)233-0644  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4707 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..4707  
 US-09-181-706-1

## Alignment Scores:

Pred. No.: 1.9e-134 Length: 4707  
 Score: 1349.00 Matches: 480  
 Percent Similarity: 39.78% Conservative: 291  
 Best Local Similarity: 24.77% Mismatches: 591  
 Query Match: 13.50% Indels: 576  
 DB: 3 Gaps: 74

US-09-964-956-13 (1-1896) x US-09-181-706-1 (1-4707)

Qy 72 TyrLysLeuSerSerAspLeuLysValLeuThrHisGluThrGlyProAspGluAsp 91  
 Db 208 TACAGCTGGAGCACAGCTCTCGCGCTGACCGGACCAAGCGGC----- 255  
 Qy 92 AsnProLysCysTyrProArgIleValGlnThrCysAsnGluProLeuThrThr 111  
 Db 256 -----AACTGCACAGAGCGCGTCTCGCTGCGC 282  
 Qy 112 -----AsnAsnValAsnLysMetLeuLeuIleAspTyrLys 123  
 Db 283 CCCCCCGCGGCGCGCGCGCGGAGCAGCTTACAGCAAG---CTGCTGCTGCCCTACCGC 339  
 Qy 124 GluAsnArg-----LeulleAlaCysGlySerLeuTyrGlnGlyIle 137  
 Db 340 GAGGGGCGCGCGCTCGGGGGGCTGCTGCTCACCGGCTGACCTTCGACCGGGGCGCC 399  
 Qy 138 CysLysLeuLeuArgLeuGluAspLeuPheLysLeuGluProTyrHisLysLysGlu 157  
 Db 400 TCGAGGTGGCGCCCTGGGCAACCTG-----AGCGGCAAC 435  
 Qy 158 HisTyrLeuSerGlyVal-----AsnGluSerGlySerValPheGlyVal 172  
 Db 436 TCCCTGCGCACGCGCACCGAGGTGCTGCTGCCACCGCGAGGCTCGACGGCGCGCTG 495  
 Qy 173 IleValSerTyrSerAsnLeuAspLysLeuPheIleAla----- 186  
 Db 496 GTGTACCGC---GCGGGCGGAACAACGCTGTGTACCTGGCGGTGGCGCGCACCTACGCTG 552  
 Qy 187 -----ThrAla 188  
 Db 553 CTGCCTGAGCGGAGAGCGGAGCGGTGCAACCCCGGCGCATCCGACACGACAGCGCC 612  
 Qy 189 ValAspGlyLysProGluTyrPheProThrIleSerArgLysLeuThrLysAsnSer 208  
 Db 613 ATCCGCTCAGGACCGGAGGGCGGCGAGCTGCCACGAGGAGCTGGGGCGGCTCAAG 672  
 Qy 209 GluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheVal---AlaSerMetIle 227  
 Db 673 CTGTGCGAGGGCGGCGGCGCTGCTGCTGCGACGCTTTCTCTGGAAGCGGACGATC 732  
 Qy 228 LysIleProSerThrPheThrIleIleProAspPheAspIleTyrTyrValTyrGly 247  
 Db -----

Db 733 TACTTCCCC-----TACTACCCTACAAC 756  
 Qy 248 PheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMetValSerProPro 267  
 Db 757 TATACGAGCGGC-----GCTGCACCGGCTGGCCAGCATGGCCGCCATCGCG 804  
 Qy 268 GlySerThrThrLysGluGlnValTyrThrSerLysLeuValArgLeuCysLysGluAsp 287  
 Db 805 CAGAGCACC-----GAG 816  
 Qy 288 ThrAlaPheAsnSerTyrValGluValProIleGlyCysGluArgSerGlyValGlu--- 306  
 Db 817 GTGCTGTTCCAGGCG-----CAGGCATCCCTCGACTGCGGCACGCCACCCCGCGCGC 870  
 Qy 307 TyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeu 326  
 Db 871 CGCCGCTGCTCTCTCTCCAGCCTAGTGGAGGCC----- 906  
 Qy 327 GlyValHisProAspAspLeuLeuPheThrValPheSer-----LysGlyGlnLys 344  
 Db 907 -----CTGCACGCTCTGGCGGGAGTCTTCAGCGCGCGCTGCAGAGGCG 951  
 Qy 345 ArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysGlnIleAsn 364  
 Db 952 CAGGAGCGCGCTCCCGCCACCCAGCGGCTCTGCTCTTTCAGAAATGAGTGAGATCCAG 1011  
 Qy 365 AspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThrLeuAspLeuAla 384  
 Db 1012 GCGCGCGCCACAGG-----GTCAGC 1032  
 Qy 385 Trp---LeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIleAspAsn 403  
 Db 1033 TGGGACTTCAAGACGCGCGAGAGCCACTCAAGAAAGGG----- 1071  
 Qy 404 PheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArgGlyIlePro 423  
 Db 1072 -----GATCAACCT-----GAAAGAGTCCCAACCAATCCCA 1101  
 Qy 424 ValPheThrGluAspArgAspMetThrSerValIleAlaTyrValTyrLysAsnHis 443  
 Db 1102 TCATCTACCTTCATCCATCCGACCTGACATCGTTTATGGCACCCTGGTAAATGAACAGG 1161  
 Qy 444 SerLeuAlaPheValGlyThrLysSerGlyLysLeuLysIleArgValAspGlyPro 463  
 Db 1162 ACTGTTTATTCTTGGGACTGGAGATGGCCAGTTACTTAAGTTATTCTT----- 1212  
 Qy 464 ArgGlyAsnAlaLeuGlnTyrGluThrValGlnValVal----- 476  
 Db 1213 ---GGTGAGAAATTGACTTCAAAATGTCAGAGGTTATCTATGAAATTAAGAGAGACA 1269  
 Qy 477 -----AspProGlyProValLeuArgAspMetAlaPheSerLysAsp 490  
 Db 1270 CCGTGTTCCTACAAACTCGTTCCTGATCCTGTG-----AAG 1305  
 Qy 491 HisGluGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCys 510  
 Db 1306 AATATCTACATTTATCTAAACAGTGGGAAAGAGGTGAGGAGAATTCGTGTGCAAACTGC 1365  
 Qy 511 GlyGlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTrpCys 530  
 Db 1366 AATAACATATAATCTGTCGAGTGTAAACAGCCACAGACCCCTACTCGGTTGGTGC 1425  
 Qy 531 ValLeuHisAsnThrCysThrArgLysGluArgCysGluArgSerLysGluProArgArg 550  
 Db 1426 CATTCGCTACAAAGGTGCACCTTTTCAAGGAGATTGT----- 1461  
 Qy 551 PheAlaSerGluMetLysGlnCysValArgLeuThrValHisProAsnAsnIleSerVal 570  
 Db 1462 -----GTACATTCAGAGAACTTAGAA--- 1482  
 Qy 571 SerGlnTyrAsnValLeuLeuValLeuGluThrTyrAsnValProGluLeuSerAlaGly 590  
 Db 1483 -----AACTGGCTGGGATATTCGCTCGGA 1506

QY 591 ValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValGlyAsnGlnIle 610  
Db 1506 -----  
QY 611 GlnCysTyrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAspHis 630  
Db 1507 -----GCAAAAGTGCCTAAATTT-----  
QY 631 HisValValGlnLeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerThrSer 650  
Db 1528 -----CAGATAATTCGACGAGTAAGAAAGACTACAGTACGATGTTGGTGGGAGC 1578  
QY 651 PheValPheTyrAsnCysSerValHisAsnSerCysLeu----SerCysValGluSerPro 669  
Db 1579 TTC-----TCTCCAAGACACTCAAGTGCATGTTGTAAGAAATGTGGACTCTAGC 1626  
QY 670 TyrArgCysHisTrpCysLysTyrArgHisValCysThrHisAspProLysThrCysSer 689  
Db 1627 -----AGGAGCTCTGC-----  
QY 690 PheGlnGluGlyArgValLysLeuProGluAspCysProGlnLeuLeuArgValAspLys 709  
Db 1638 -----  
QY 710 IleLeuValProValGluValIleLysProIleThrLeuLysAlaLysAsnLeuProGln 729  
Db 1639 -----CAGATAAAAGTCAG 1653  
QY 730 ProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGluGlnArg 749  
Db 1654 CCAACCGG-----ACCTGCACCTGTAGCATC-----  
QY 750 ValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThrSerTyrSerTyr 769  
Db 1681 -----CCACACGACGACCTACAAGATGTTTCAGTTGTCAACGTCGATGTTCTCTTC 1734  
QY 770 GluGlyMetGluIleAsnAsnLeuProValGluLeuThrValValTyrAsn-----Gly 787  
Db 1735 GGTCTCT-----TGGAAATTTATCAGAC 1755  
QY 788 HisPheAsnIleAspAsnProAlaGlnAsnLysValHisLeuTyrLysCysGlyAlaMet 807  
Db 1756 AGATTCAACTTACCAAC-----TGCTCATCATTA 1785  
QY 808 ArgGluSerCysGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTrpCysGln 827  
Db 1786 AAAGAA---TGCCACGATCGGTAGAACT-----GGCTGGCGGTGGTAA 1830  
QY 828 GlyProGlyGlnCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuGlu 847  
Db 1831 AGTGCAAGAAGGTGT----- 1845  
QY 848 LeuSerGlyAlaLysSerLysCysThrAsnProArgIleThrGluIleIleProValThr 867  
Db 1846 -----ATCCACCCCTTCACA 1860  
QY 868 GlyProArgGluGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPhe 887  
Db 1861 GCT----- 1863  
QY 888 ArgAspIleAlaSerHisValLysValAlaGlyValGluCysSerProLeuValAspGly 907  
Db 1864 -----TCCGACCCCT-----TCT 1875  
QY 908 TyrlleProAlaGluGlnIleValCysGluMet-----GlyGluAla 921  
Db 1876 GATTATGAGAAACACGAGAACAGTGTCCAGTGTGTCGAGAGACATCAGGAGGAGGA 1935  
QY 922 LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgProGluPhe 941  
Db 1936 AGACCCAAAGAGAACAGGG-----AACAGA 1962

QY 942 MetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysPro 961  
Db 1963 ACCAACCAAGCTTTACAGGTCTCTAC-----ATTAAGTCCATTGAGCCA 2007  
QY 962 SerArgGlyProMetSerGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAla 981  
Db 2008 CAGAAAGTATCGACATTAGGGAAACAGCGTAGTAGTACGCGGAGCAAACTTTACCCGG 2067  
QY 982 GlySerAsnVal---ValValMetPheGlyLysGlnProCys----- 994  
Db 2068 GCATCGAACATCACAATGATCTGAAAGGACCACTACCTGTGATAGGATGTGATACAG 2127  
QY 995 LeuPheHisArgArgSerProSerTyrIleValCysAsnThrThrSerSerAspGluVal 1014  
Db 2128 GTTAGCCCATGTCTAAATGACACCCACATGAAATTTCTCTTCCATCAGCCGGA--- 2184  
QY 1015 LeuGluMetLys---ValSerValGlnValAspArgAlaLysIleHisGlnAspLeuVal 1033  
Db 2185 ---GAAATGAAGGATGTGTATCCAGTTGTAGTGGTGGGAACGTCTTCTGTGGGATCC 2241  
QY 1034 PheGlnTyrValGluAspProThrIleValArgIleGluProGluTrpSerIleValSer 1053  
Db 2242 TTATCTACATTGCTCTGCCACATGTTCCCTTATATTTCTCTGCTACCACCTGGATCAGT 2301  
QY 1054 GlyAsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIle 1073  
Db 2302 GGTGTCAAAATATAACCATGATGGCAGAAATTTTGTATGATGTAATTGACAAC-----TTA 2355  
QY 1074 ArgAlaLysHisGlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGlu 1093  
Db 2356 ATCATTTTCATGAAATTAAGGAAACATAAATGCTCTGAAATATTGTGTGGCGACTTAC 2415  
QY 1094 MetThrCysGlnAlaProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGlu 1113  
Db 2416 TCCGGGTTTTAGCCCCCAGTTTA----- 2439  
QY 1114 ArgProGluGluPheGlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeuAsnLys 1133  
Db 2440 -----AAGAGTTCAAAAGTCGCAGCAGATGTCACGTGTAAGCTGAGAGTACAAGAC 2490  
QY 1134 Thr-----AsnPheThrTyrTyrProAsnProValPheGluAlaPheGly 1148  
Db 2491 ACCTACTTGGATTGTGGAAACCTCGCATGTCGGGAGGAGCCAGATTTCACGGGTATPCGG 2550  
QY 1149 ProSerGlyIleLeuGluLeuLysProGlyThrProIleIleLeuLysGlyLysAsnLeu 1168  
Db 2551 -----GTGCAATCCGAGGTGGACACAGAACTGCAAAATTCAA----- 2592  
QY 1169 IleProProValAlaGlyCysValLysLeuAsnTyrThrValLeuValGlyGluLys 1188  
Db 2593 -----AAAGAAATGACAACTTCAATATTTCCAAAAA 2625  
QY 1189 ProCysThrValThrVal-----SerAspValGlnLeuLeuCysGluSerProAsp 1205  
Db 2626 GACATTGAAATTTACTCTCTTCCATGGGAAATGGGCAATTAATTTGCACTTTTGAAT 2685  
QY 1206 LeuIle----- 1207  
Db 2686 ATTACTAGAAATCAAGATCTTACCACCATCTTTGCCAAATTAAGGCATCAAGACTGCA 2745  
QY 1208 -----GlyArgHisLysValMetAlaArgValGlyGlyMetGluTyrSer 1222  
Db 2746 AGCACCAATTGCCAACTCTTCTAAGAAAGTTCGGGTCAAGCTGGGAAACCTGGAG----- 2799  
QY 1223 ProGlyMetValTyrIleAlaProAspSer---ProLeuSerLeuProAlaIleValSer 1241  
Db 2800 -----CTACGTCGAGCAGGAGTCAGTTCTTCCACATGATGATTTTGTATGTGTC 2850  
QY 1242 IleAlaValAlaGlyGlyLeuIleIlePheIleValIleAlaValLeuIleAlaTyrLys 1261  
Db 2851 CTCCTCTGTC-----TTGCTAGTAGTATGTCATTTTCCGCGCCGTGGGGTGACAGG 2901  
QY 1262 ArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeu 1281

```
Db 2902 CACAAATCGAAGGAG-----CTGAGTCGCAACAGAGCTCAA--CAACTAGAATTGCTG 2952
Qy 1282 GluSerArgValAlaLeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIleHis 1301
Db 2953 GAAAGCGAGCTCCGCAAGAGATACGTCAGCGCTTTGCTGAGCTGCACATGATGAATG 3012
Qy 1302 GluLeuThrSerAspLeuAspGlyAlaGly---IleProPheLeuAspTyrArgThrTyr 1320
Db 3013 GATGTG-----GTTGATAGTTTGGAACTGTTCCTTCTTGTACTACAAACATTTT 3063
Qy 1321 ThrMetArgValLeuPro-----GlyIleGluAspHisProValLeuArgAspLeu 1338
Db 3064 GCTGTGAGAACTTCTTCCCTGAGCTCAGTGGCTTCCACCATCTTCTCACTGAAGATATG 3123
Qy 1339 GluValProGlyTyrArgGlnGluArgValGluLysGlyLeuLeuPheAla----- 1356
Db 3124 -----CATACAGAGAGCCACACAGCAAGTAAGTCTCACAGCTTTGGAT 3171
Qy 1357 GlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArg 1376
Db 3172 GCCCTAATCTGTAATAAAGCTTCTTGTACTGTCTCATCCACACCTTGAAGAGCAGAAG 3231
Qy 1377 SerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSer 1396
Db 3232 AACTTTTCTGTGAAGGACAGTGTCTGTTGCCCTTCTTAACCATTTGCACCTGCAAAACC 3291
Qy 1397 LysLeuGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsn 1416
Db 3292 AAGCTGGTCTACCTGACCAAGCATCTAGAGGTGCTGACCAAGGACTTGTATGGAACAGTGT 3351
Qy 1417 LeuGluSerLysAsnHisProLysLeuLeuLeuArgThrGluSerValAlaGluLys 1436
Db 3352 -----AGTAACATGCAAGCCGAAATCATGTGAGACGACGAGGTCGTCGTCGAAAAA 3405
Qy 1437 MetLeuThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlu 1456
Db 3406 CTCCTCAAACTGGATGCTGCTGCTTCTGGAATTTCTCCGGGAGACTGCTCGGAGAG 3465
Qy 1457 ProLeuPheSerLeuPheCysAlaIleLysGlnMetGluLysGlyProIleAspAla 1476
Db 3466 CCTTCTATTGCTGGTGACGACTCTGAACACAGAAAAATTAACAGGGTCCCGTGTGATGTA 3525
Qy 1477 IleThrGlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnIleAsp 1496
Db 3526 ATCACTTGCAAAGCCCTGTACACACTTAATGAAGACTGGCTGTTGGCAGGTTCCGGAA 3585
Qy 1497 TyrLysThrLeuValLeuSerCysVal-----SerProAspAsnAlaAsnSerPro 1513
Db 3586 TTCAGTACTGGCATTAACGTCGCTTTGAAAAAATCCCGAAAAACGAGAGTGCAGAT 3645
Qy 1514 -----GluValProValLysIleLeuAsnCysAspThrIleThrGlnValLysGlu 1530
Db 3646 GTCTGCTGGAATATTTCAGTCAATGTTCTGACTGTGACCATTTGGCCAGCAAGCAAGAA 3705
Qy 1531 LysIleLeuAspAlaIlePheLysAsnValProCysSerHisArgProLysAlaAlaAsp 1550
Db 3706 AAGATTTCGAAGCATTTTAAAGCAAAATGGCTCTCCTTATGGACTTCAGCTTAATGAA 3765
Qy 1551 MetAspLeuGluTrpArgGlnGlySerGlyAlaArgMetIleLeuGlnAspGluAspIle 1570
Db 3766 ATTGCTCTTGAGCTTCAATGGGCACACGACAGAAAGAACITCTTGACATCGACATTC 3825
Qy 1571 ThrThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHisTyrGlnValPro 1590
Db 3826 TCCGTGATTCTGAAGATGGAATCAACCAAGCTAAACACCATTTGGCCACTATGAGATATCA 3885
Qy 1591 AspGlySerValValAlaLeuValSerLysGlnValThrAlaTyrAsnAlaValAsnAsn 1610
Db 3886 AATGGATCCACTAAAAAGTC----- 3906
Qy 1611 SerThrValSerArgThrSerAlaSerLysTyrGluAsnMetIleArgTyrThrGlySer 1630
```

```
Db 3907 -----TTTAAGAAGATAGACAATTTTACTTCAGAT 3936
Qy 1631 ProAspSerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyValLys 1650
Db 3937 GTGGAGTACTCGGATCACCACATCCCATTTGATTTTACCAGATTCGGAAGCA----- 3987
Qy 1651 MetTrpHisLeuValLysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGlySer 1670
Db 3988 ---TTCAGAGATGTGCAAGGAAAGAGACAT-----CGAGGGAAG 4023
Qy 1671 Lys-----MetValSerGluIleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeu 1688
Db 4024 CACAGTTCAAAGTAAAGAAATGATCTGACAAAGCTGCTGCGACCAAGGTGGCAATT 4093
Qy 1689 GlnLysPheValAspAspLeuPheGluThrIlePheSerThrAlaHisArgGlySerAla 1708
Db 4084 CATTCGTGCTTGAAAAACTTTTACAGACATTTGGAGTTTACCCAAC-----AGCAGA 4137
Qy 1709 LeuProLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGly 1728
Db 4138 GCTCCATTTGCTATAAAATACTTTTGTGACCTTTTGGACCCCGAGCTGAAAAACAAAAA 4197
Qy 1729 IleHisAspProHisValArgHisThrTrpLysSerAsnCysLeuProLeuArgPheTrp 1748
Db 4198 ATCAGATCCTGACGCTGCTACATATTTGGAAACAAACACACCTTCTCTTCGCTCTGG 4257
Qy 1749 ValAsnMetIleLysAsnProGlnPheValPheAspIleHisLysAsnSerIleThrAsp 1768
Db 4258 GTAAACATCCTGAAGAACCTCGATGTTGCTTTGACATTAAGAAGACACACCATATAGAC 4317
Qy 1769 AlaCysLeuSerValValAlaGlnThrPheMetAspSerCysSerThrSerGluHisArg 1788
Db 4318 GCCTGTTGCTGAGTGTGCTCCAGGATTCATGATGATGCTTCTCTCAGACAGCAGCAA 4377
Qy 1789 LeuGlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysAspIleProSerTyrLys 1808
Db 4378 CTAGGAAGGAAGCACCACCACTATAGCTTCTCTATGCCAAGATATCCCAACCTACAA 4437
Qy 1809 AsnTrpValGluArgTyrTyrSerAspIleGlyLysMetProAlaIleSerAspGlnAsp 1828
Db 4438 GAAGAGTAAATCTTTATTACAAAGCAATCAGGATTTGCTCATTTGTCATCCTCAGAA 4497
Qy 1829 MetAsnAlaTyrLeuAlaGluGlnSerArgMetHisMetAsnGluPheAsnThrMetSer 1848
Db 4498 ATGGAAGAAATTTTAAGTCTGAGGATCTAAGAAACATGAATTAATTAAGAGAGTG 4557
Qy 1849 AlaLeuSerGluIlePheSerTyrValGlyLysTyrSerGluGluIleLeuGlyProLeu 1868
Db 4558 GCCTTCAGCAAAATTTTACAAATACATCGTAAAAATATTTTGTATGAGATCTAAATAACTA 4617
Qy 1869 AspHisAsp-----AspGlnCysGlyLysGlnLysLeuAlaTyrLysLeu 1883
Db 4618 GAAAGAGACGAGGGCTGGAAGAGCTCAGAAACAACTCTTTCATGTAAAGTC 4671
```

## RESULT 2

```
US-09-458-791-1
; Sequence 1, Application US/09458791
; Patent No. 6174689
; GENERAL INFORMATION:
; APPLICANT: Spriggs, Melanie
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; RECEPTOR DNA AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: MS-DOS/Windows 95  
SOFTWARE: Word for Windows 95, 7.0a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/458,791  
FILING DATE: 10-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/958,598  
FILING DATE: 28-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Henry, Janis C  
REGISTRATION NUMBER: 34,347  
REFERENCE/DOCKET NUMBER: 2631  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)470-4189  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4707 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4707  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-458-791-1

Alignment Scores:  
Pred. No.: 1,9e-134 Length: 4707  
Score: 1349.00 Matches: 480  
Percent Similarity: 39.78% Conservative: 291  
Best Local Similarity: 24.77% Mismatches: 591  
Query Match: 13.50% Indels: 576  
DB: 4 Gaps: 74

US-09-964-956-13 (1-1896) x US-09-458-791-1 (1-4707)

QY 72 TyrLysLeuSerSerAspLeuLysValLeuValThrHisGluThrGlyProAspGluAsp 91  
DB 208 TACAGCTGAGCAGACAGCCTCGCGCCCTGTACCGGACCAACGCGGC----- 255  
QY 92 AsnProLysCysTyrProProArgIleValGlnThrCysAsnGluProLeuThrThrThr 111  
DB 256 -----AACTGCACAGCGCGTCTCGCTGGCG 282  
QY 112 -----AsnAsnValAsnLysMetLeuLeuLeuAspTyrLys 123  
DB 283 CCCCCCGCGCGCCCGCGCGGAGCAGCTTCAGCAAG---CTGCTGTGCTACCGC 339  
QY 124 GluAsnArg-----LeuIleAlaCysGlySerLeuTyrGlnGlyIle 137  
DB 340 GAGGGGGCGCGCGCTCGGGGGCTGCTGCACCGCTGACCTCGACCGGGGCGCC 399  
QY 138 CysLysLeuLeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLysGlu 157  
DB 400 TCGAGGTGCGCCCTCGGCAACCTG-----ACCGCGCAAC 435  
QY 158 HisTyrLeuSerGlyVal-----AsnGluSerGlySerValPheGlyVal 172  
DB 436 TCCCTCGCAACGGCACCGAGGTGTGTGCTGCGCACCGCGAGGCTCGACGGCGCGGTG 495  
QY 173 IleValSerTyrSerAsnLeuAspLysLeuPheIleAla----- 186  
DB 496 GTGTACCGC---CGGGCGGAAACACCGCTGTGTACCTGCGGTGGCGCCACCTACGTG 552  
QY 187 -----ThrAla 188  
DB 553 CTGCCTGAGCGGAGCGGCGCGCTGCACCCCGCGGCATCCGACACGACGCGCC 612

QY 189 ValAspGlyLysProGluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSer 208  
DB 613 ATCCGCGTCAAGGACACGAGGCGCGAGCTGGCCACGAGGAGCTGGGGCGCTCAAG 672  
QY 209 GluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheVal---AlaSerMetIle 227  
DB 673 CTGTGCGAGGCGCGGCGAGCTGCTGCTGCGAGCGCTTCTCTGGAACGCGCATC 732  
QY 228 LysIleProSerAspThrPheThrIleIleProAspPheAspIleTyrValTyrGly 247  
DB 733 TACTTCCCG-----TACTACCCCTACAAC 756  
QY 248 PheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMetValSerProPro 267  
DB 757 TATACGAGCGCG-----GCTGCCACCGGCTGGCCGAGCATGCGCGCATCGCG 804  
QY 268 GlySerThrThrLysGluGlnValTyrThrSerLysLeuValArgLeuCysLysGluAsp 287  
DB 805 CAGAGCAC-----GAG 816  
QY 288 ThrAlaPheAsnSerTyrValGluValProIleGlyCysGluArgSerGlyValGlu--- 306  
DB 817 GTGCTGTCCAGGCG-----CAGGCATCCCTGAGCTGGCGCCACGCGCCACCGCGC 870  
QY 307 TyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeu 326  
DB 871 CGCGCGCTGCTCTCTCTCCAGCTAGTGAGGCG----- 906  
QY 327 GlyValHisProAspAspLeuLeuPheThrValPheSer-----LysGlyGlnLys 344  
DB 907 -----CTGGAGCTGTGGCGGAGGTTCAGCGCGCGCTGGAGGCGC 951  
QY 345 ArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysGlnIleAsn 364  
DB 952 CAGGAGCGCGCTCCCGCCACCCAGCGCGCTGCTCTCTCAGATGAGTGAGATCCAG 1011  
QY 365 AspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThrLeuAspLeuAla 384  
DB 1012 GCGCGCGCCAAAGAG-----GTACGC 1032  
QY 385 Trp---LeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIleAspAspAsn 403  
DB 1033 TGGAGCTTCAAGCGCGGAGAGGCGCTGCAAGAGAGG----- 1071  
QY 404 PheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArgGlyIlePro 423  
DB 1072 -----GATCAACCT-----GAAAGAGTCCCAACCAATCGCA 1101  
QY 424 ValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrValTyrLysAsnHis 443  
DB 1102 TCATCTACCTTGATCCCATTCGACCTGACATCCGTTTATGGCAGCGTGTGTAATGAACAG 1161  
QY 444 SerLeuAlaPheValGlyThrLysSerGlyLysLeuLysLysIleArgValAspGlyPro 463  
DB 1162 ACTGTTTATCTTGGGACCTGGAGATGCGAGTTACTTAAGGTATTCTT----- 1212  
QY 464 ArgGlyAsnAlaLeuGlnTyrGluThrValGlnValVal----- 476  
DB 1213 ---GGTCAGAAATTTGACTTCAAAATTTGTCAGAGGTTATCTATGAAATTAAGAAGAGACA 1269  
QY 477 -----AspProGlyProValLeuArgAspMetAlaPheSerLysAsp 490  
DB 1270 CCTCTTTTCTACAACTCGTTCTCTGATCTCTG-----AAG 1305  
QY 491 HisGluGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCys 510  
DB 1306 ANTATCTACATTTATCTAACAGCTGGGAAAGAGGTGAGAGAAATTCGTGTTCACAACCTGC 1365  
QY 511 GlyGlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTyrCys 530  
DB 1366 AATAACATAAATCTCTGTCGAGTGTTTACAGCCACGAGACCTCACTCGCGGTGGTGC 1425  
QY 531 ValLeuHisAsnThrCysThrArgLysGluArgCysGluArgSerLysGluProArgArg 550

Db 1426 CATTGCGTACAAAGGTGCACTTTTCAAGGAGATTGT----- 1461  
QY 551 PheAlaSerGluMetLysGlnCysValArgLeuThrValHisProAsnAsnIleSerVal 570  
Db 1462 -----GTACATTCCAGAGAAGCTTAGAA--- 1482  
QY 571 SerGlnTyrAsnValLeuLeuValLeuGluThrTyrAsnValProGluLeuSerAlaGly 590  
Db 1483 -----AAGTGGCTGGATATTTCTGCTGGA 1506  
QY 591 ValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValGlyAsnGlnIle 610  
Db 1506 ----- 1506  
QY 611 GlnCysTyrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAspHis 630  
Db 1507 -----GCAAAAGTGCCTAAAAAT 1527  
QY 631 HisValValGlnLeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerThrSer 650  
Db 1528 -----CAGATAATTCGAGCAGTAAGAAAGACTACAGTGACTATGGTGGGAGC 1578  
QY 651 PheValPheTyrAsnCysSerValHisAsnSerCysLeu---SerCysValGluSerPro 669  
Db 1579 TTC-----TCTCCAAGACACTCAAGTGCATGGTGAAGAATGTGGACTCTAGC 1626  
QY 670 TyrArgCysHisTrpCysLysTyrArgHisValCysThrHisAspProLysThrCysSer 689  
Db 1627 -----AGGAGCTCTGC----- 1638  
QY 690 PheGlnGluGlyArgValLysLeuProGluAspCysProGlnLeuLeuArgValAspLys 709  
Db 1638 ----- 1638  
QY 710 IleLeuValProValGluValIleLysProIleThrLeuLysAlaLysAsnLeuProGln 729  
Db 1639 -----CAGAATAAAGTCAG 1653  
QY 730 ProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGluGlnArg 749  
Db 1654 CCCAACCG-----ACCTGCACCTGTAGCATC----- 1680  
QY 750 ValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThrSerTyrSerTyr 769  
Db 1681 -----CCAACGACAGAACCTACAAAGATGTTTCAGTTGTCAACGTGATGTTCTCCTC 1734  
QY 770 GluGlyMetGluIleAsnAsnLeuProValGluLeuThrValValTrpAsn-----Gly 787  
Db 1735 GGTCT-----TGGAATTTATCAGAC 1755  
QY 788 HisPheAsnIleAspAsnProAlaGlnAsnLysValHisLeuTyrLysCysGlyAlaMet 807  
Db 1756 AGATTCAACTTTACCAAC-----TGCTCATCAATTA 1785  
QY 808 ArgGluSerCysGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTrpCysGln 827  
Db 1786 AAAGAA---TGCCGAGCATGCGTAGAACT-----GGCTGCGCGTGGGTGATAA 1830  
QY 828 GlyProGlyGlnCysThrLeuArgGlnHisCysProAlaGlnGlnLysThrTrpLeuGlu 847  
Db 1831 ACTGCAAGAAGTGT----- 1845  
QY 848 LeuSerGlyAlaLysSerLysCysThrAsnProArgIleThrGluIleIleProValThr 867  
Db 1846 -----ATCCACCCCTTCACA 1860  
QY 868 GlyProArgGluGlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPhe 887  
Db 1861 GCT----- 1863  
QY 888 ArgAspIleAlaSerHisValLysValAlaGlyValGluCysSerProLeuValAspGly 907

Db 1864 -----TGCACCCCT-----TCT 1875  
QY 908 TyrIleProAlaGluGlnIleValCysGluMet-----GlyGluAla 921  
Db 1876 GATTATGAGAAACCAGCAACAGTGTCCAGTGGCTGTCGAGAAGACATCAGGAGGGA 1935  
QY 922 LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgProGluPhe 941  
Db 1936 AGACCAAGGAGAACCAAGGG-----AACAGA 1962  
QY 942 MetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysPro 961  
Db 1963 ACCAACAGGCTTTACAGGTCTTCAC-----ATTAAAGTCCATTGAGCCA 2007  
QY 962 SerArgGlyProMetSerGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAla 981  
Db 2008 CAGAAAGTATCGACATTAGGGAAGCAACGATAGTAGTAACGGAGCAAACTTTACCCTCG 2067  
QY 982 GlySerAsnVal---ValValMetPheGlyLysGlnProCys----- 994  
Db 2068 GCATCGAACATCACAATGATCTGAAAGGAACACGATACCTGTCTAAGGATGTGATACAG 2127  
QY 995 LeuPheHisArgArgSerProSerTyrIleValCysAsnThrThrSerSerAspGluVal 1014  
Db 2128 GTTACCATGTCTAAATGACACCCACATGAAATCTCTCTCCATCAAGCCGGAAA--- 2184  
QY 1015 LeuGluMetLys---ValSerValGlnValAspArgAlaLysIleHisGlnAspLeuVal 1033  
Db 2185 ---GAATGACAGGATGTGTATCCAGTTTGTGTTGGGAACGTCTTCTGTGGGATCC 2241  
QY 1034 PheGlnTyrValGluAspProThrIleValArgIleGluProGluTrpSerIleValSer 1053  
Db 2242 TTATCCTACATTGCTCGCCACATGTTCCCTTATATTTCCCTGCCCTCCAGCTGATCAGT 2301  
QY 1054 GlyAsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIle 1073  
Db 2302 GGTGTCAAATATATACCATGATGGCAGAAAATTTTGATGTAATTCACACAC-----TTA 2355  
QY 1074 ArgAlaLysHisGlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGlu 1093  
Db 2356 ATCATTTTCATGAATTAAGAAAGAAACATAAATGCTCTGAATATTGTGTGGCGACTTAC 2415  
QY 1094 MetThrCysGlnAlaProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGlu 1113  
Db 2416 TCGGGGTTTTAGCCCCCAGTTTA----- 2439  
QY 1114 ArgProGluGluPheGlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeuAsnLys 1133  
Db 2440 -----AAGAGTTCAAAAGTGGCCACCAATGTCTACTGTGAAGCTGACAGTACAAGAC 2490  
QY 1134 Thr-----AsnPheThrTyrTyrProAsnProValPheGluAlaPheGly 1148  
Db 2491 ACCTACTTGGATTGTGGAACCCCTGCAGTATCGGGAGGACCCAGATTTCACGGGGTATCGG 2550  
QY 1149 ProSerGlyIleLeuGluLeuLysProGlyThrProIleIleLeuLysGlyLysAsnLeu 1168  
Db 2551 -----GTGGAATCCGAGGTGGACACAGAACTGGAAGTGAATAATTCAA----- 2592  
QY 1169 IleProProValAlaGlyGlyAsnValLysLeuAsnTyrThrValLeuValGlyGluLys 1188  
Db 2593 -----AAAGAAATGACAACTTCAATATTTCCAAAAAA 2625  
QY 1189 ProCysThrValThrVal-----SerAspValGlnLeuLeuCysGluSerProAsn 1205  
Db 2626 GACATTGAAATTAATCTCTTCCATGGGAAAATGGGCAATTAATTCAGTGTTCGAAAT 2695  
QY 1206 LeuIle----- 1207  
Db 2686 ATTACTAGAATCAAGATCTTACCACCATCTTTGCCAAAATTAAGGCATCAAGACTGCA 2745  
QY 1208 -----GlyArgHisLysValMetAlaArgValGlyGlyMetGluTyrSer 1222  
Db 2746 AGCACCATTCGCCAACTCTTTTAAGAAAGTTCGGGTCAAGCTGGGAAACCTGGAG----- 2799



;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Janis C. Henry  
;; STREET: 51 University St.  
;; CITY: Seattle  
;; STATE: WA  
;; COUNTRY: US  
;; ZIP: 98101  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: MS-DOS/Windows 95  
;; SOFTWARE: Word for Windows 95, 7.0a  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/459,066  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/958,598  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Henry, Janis C  
;; REGISTRATION NUMBER: 34,347  
;; REFERENCE/DOCKET NUMBER: 2631  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206)470-4189  
;; TELEFAX: (206)233-0644  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4707 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; HYPOTHEtical: NO  
;; ANTI-SENSE: NO  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..4707  
;; US-09-459-066-1

Alignment Scores:  
Pred. No.: 1.9e-134 Length: 4707  
Score: 1349.00 Matches: 480  
Percent Similarity: 39.78% Conservative: 291  
Best Local Similarity: 24.77% Mismatches: 591  
Query Match: 13.50% Indels: 576  
DB: 4 Gaps: 74

US-09-964-956-13 (1-1896) x US-09-459-066-1 (1-4707)

QY 72 TyrLysLeuSerSerAspLeuLysValLeuValThrHisGluThrGlyProAspGluAsp 91  
Db 208 TACAGCTGGAGCACACCTCTCGCGCTGTACCGGGACCAAGCGGC----- 255  
QY 92 AsnProLysCysTyrProProArgIleValGlnThrCysAsnGluProLeuThrThr 111  
Db 256 -----AACTGCACAGACCGGTCTCGCTGGCG 282  
QY 112 -----AsnAsnValAsnLysMetLeuLeuIleAspTyrLys 123  
Db 283 CCCCCCGGGCCCCCGCGCGGAGCAGCTTCAGCAG-----CTGCTGCTCCCTACCGC 339  
QY 124 GluAsnArg-----LeullealaCysGlySerLeuTyrGlnGlyIle 137  
Db 340 GAGGGGGCGGGCGCTCGGGGGGCTGTCTGCTCACCGGCTGGACCTTCGACCGGGCGCC 399  
QY 138 CysLysLeuLeuArgLeuGluAspLeuPhelLysLeuGluProTyrHisLysLysGlu 157  
Db 400 TGCAGAGTGGCGCCCTGGGCAACCTG-----AGCGCCAC 435  
QY 158 HisTyrLeuSerGlyVal-----AsnGluSerGlySerValPheGlyVal 172  
Db 436 TCCCTGGCCACAGCACCGAGGTGTGTGTCGTGCCACCGCGAGGCTCGACGCGCGCGTG 495

QY 173 IleValSerTyrSerAsnLeuAspAspLysLeuPheIleAla----- 186  
Db 496 GTGTACCGC---CGGGCGCGAACAACCGCTGGTACCTGGCGGTGGCGCGCACCTACGTG 552  
QY 187 -----ThrAla 188  
Db 553 CTGCCTAGCCGGAGACGGCGGCTGCAACCCCGGGCATCCGACCCAGACACGCGC 612  
QY 189 ValAspGlyLysProGluTyrPheProThrIleSerArgLysLeuThrLysAsnSer 208  
Db 613 ATCGCGCTCAAGGACACGAGGGCGCGCTGCCACGACGAGGAGCTGGGGCGCTCAAG 672  
QY 209 GluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheVal---AlaSerMetIle 227  
Db 673 CTGTGCGAGGGCGCGCGCTGCACTTCGTGGACGCTTTCTCTGGAACGGCAGCATC 732  
QY 228 LysIleProSerAspThrPheThrIleIleProAspPheAspIleTyrValTyrGly 247  
Db 733 TACTTCCCC-----TACTACCCCTACAC 756  
QY 248 PheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMetValSerPro 267  
Db 757 TATACGAGCGC-----GCTGCCACCGGCTGCCACGATGCGCGCATCGCG 804  
QY 268 GlySerThrThrLysGluGlnValTyrThrSerLysLeuValArgLeuLysGluAsp 287  
Db 805 CAGAGCAC-----GAG 816  
QY 288 ThrAlaPheAsnSerTyrValGluValProIleGlyCysGluArgSerGlyValGlu 306  
Db 817 GTGCTGTTCACGGC-----CAGCATCCCTCGACTCGGGCCACCGCCGAGCGC 870  
QY 307 TyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeu 326  
Db 871 CGCGCGCTGCTCTCTCTCCAGCTAGTGGAGGC----- 905  
QY 327 GlyValHisProAspAspLeuLeuPheThrValPheSer-----LysGlyGlnLys 344  
Db 907 -----CTGGACGCTGCGGGGAGGTTCAGCGCGCGCCCTGGAGAGGC 951  
QY 345 ArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysGlnIleAsn 364  
Db 952 CAGGAGCGCGCTCCCCACCCACCGCGCTCTGCTCTCAGAATAGTGTAGATCCAG 1011  
QY 365 AspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGlyThrLeuAspLeuAla 384  
Db 1012 GCGCGCGCCCAAGAG-----GTCAGC 1032  
QY 385 Trp---LeuLysValLysAspIleProCysSerSerAlaLeuLeuThrIleAspAspAsn 403  
Db 1033 TGGGACTTCAGACGCGCGAGACCCACTGCCAAGAGG----- 1071  
QY 404 PheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetValArgGlyIlePro 423  
Db 1072 -----GATCAACCT-----GAAAGAGTCCAACCAATCGCA 1101  
QY 424 ValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrLysAsnHis 443  
Db 1102 TCATCTACCTTCATCCATTCGACCTCGACATCGTTATGCGCACCGCTGTAATGACAGG 1161  
QY 444 SerLeuAlaPheValGlyThrLysSerGlyLysLeuLysIleArgValAspGlyPro 463  
Db 1162 ACTGTTTATTCTTGGGACTGGAGATGGCAGTTACTTAAGTTATTCTT----- 1212  
QY 464 ArgGlyAsnAlaLeuGlnTyrGluThrValGlnValVal----- 476  
Db 1213 ---GGTGAGATTGTGACTTCAAAATTCAGAGAGTTATCTATCAAAATTAAGAGAGACA 1269  
QY 477 -----AspProGlyProValIleLeuArgAspMetAlaPheSerLysAsp 490  
Db 1270 CCTGTTTTCTACAACCTGCTTCCTGATCCTGTG-----AAG 1305

QY 491 HisGluGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCys 510  
Db 1306 AATATCTACATTTATCAACAGCTGGGAAAGAGGTGAGCAAAATTCGGTTGCAAACTGC 1365  
QY 511 GlyGlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTrpCys 530  
Db 1366 AATAAACAATAATCCTGTCGGAGGTGTTAAACAGCCACAGACCCCTCACTGGCGTTGGTGC 1425  
QY 531 ValLeuHisAsnThrCysThrArgLysGluArgCysGluArgSerLysGluProArgArg 550  
Db 1426 CATTCGCTACAAAGGTGCACCTTTCAAGGATGCT 1461  
QY 551 PheAlaSerGluMetLysGlnCysValArgLeuThrValHisProAsnAsnIleSerVal 570  
Db 1462 -----GTACATTGAGAACTTAGAA--- 1482  
QY 571 SerGlnTyrAsnValLeuLeuValLeuGluThrTyrAsnValProGluLeuSerAlaGly 590  
Db 1483 -----AACTGGCTGGATATTTCGTCTGGA 1506  
QY 591 ValAsnCysThrPheGluAspLeuSerGluMetAspGlyLeuValValGlyAsnGlnIle 610  
Db 1506 ----- 1506  
QY 611 GlnCysTyrSerProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAspHis 630  
Db 1507 -----GCAAAAAGTGCCTAAAT 1527  
QY 631 HisValValGlnLeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerThrSer 650  
Db 1528 -----CAGATAATTCGAAGCAGTAAAGAAAGACTACAGTGACTATGGTGGGAAGC 1578  
QY 651 PheValPheTyrAsnCysSerValHisAsnSerCysLeu---SerCysValGluSerPro 669  
Db 1579 TTC-----TCTCCAAGACACTCAAGTGCATGGTCAAGATGTGCACTCTAGC 1626  
QY 670 TyrArgCysHisTrpCysLysTyrArgHisValCysThrHisAspProLysThrCysSer 689  
Db 1627 -----AGGAGCTCTGC 1638  
QY 690 PheGlnGluGlyArgValLysLeuProGluAspCysProGlnLeuLeuArgValAspLys 709  
Db 1638 ----- 1638  
QY 710 IleLeuValProValGluValIleLysProIleThrLeuLysAlaLysAsnLeuProGln 729  
Db 1639 -----CAGAATAAAAGTCAG 1653  
QY 730 ProGlnSerGlyGlnArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGluGlnArg 749  
Db 1654 CCCAACCGG-----ACCTGCACCTGTAGCATC----- 1680  
QY 750 ValProAlaLeuArgPheAsnSerSerValGlnCysGlnAsnThrSerTyrSerTyr 769  
Db 1681 -----CCAACGAGACCACTCAACAGATGTTTCAGTTGTCACAGTGTCTCTCTTC 1734  
QY 770 GluGlyMetGluIleAsnAsnLeuProValGluLeuThrValValTrpAsn-----Gly 787  
Db 1735 GGTTCCT-----TGGAATTTATCAGAC 1755  
QY 788 HisPheAsnIleAspAsnProAlaGlnAsnLysValHisLeuTyrLysCysGlyAlaMet 807  
Db 1756 AGATTCAACTTTACCAAC-----TGCTCATCATTA 1785  
QY 808 ArgGluSerCysGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTrpCysGln 827  
Db 1786 AAAGAA-----TGCCAGCATGCGTAGAACT-----GGCTGCGCGTGGGTGATAA 1830  
QY 828 GlyProGlyGlnCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuGlu 847  
Db 1831 AGTCAAGAAGGTCT----- 1845  
QY 848 LeuSerGlyAlaLysSerLysCysThrAsnProArgIleThrGluIleIleProValThr 867

Db 1846 ----- 1846  
QY 868 GlyProArgGluGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPhe 887  
Db 1861 GCT----- 1863  
QY 888 ArgAspIleAlaSerHisValLysValAlaGlyValGluCysSerProLeuValAspGly 907  
Db 1864 -----TGCAGCCCT-----TCT 1875  
QY 908 TyrIleProAlaGluGlnIleValCysGluMet-----GlyGluAla 921  
Db 1876 GATTATGAGAGAACCAGGAACAGTGTCCAGTGTCTCGAGAAGACATCAGGAGGAGGA 1935  
QY 922 LysProSerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgProGluPhe 941  
Db 1936 AGACCCCAAGGAACAAGGGG-----AACAGA 1962  
QY 942 MetAlaArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysPro 961  
Db 1963 ACCAACCGCTTTTACAGGTCCTCTAC-----ATTAAGTCCATTGAGCCA 2007  
QY 962 SerArgGlyProMetSerGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAla 981  
Db 2008 CAGAAAGTATCGACATTAGGGAAAAGCAACGTGATAGTAAACGGGAGCAAACTTTACCCGG 2067  
QY 982 GlySerAsnVal---ValValMetPheGlyLysGlnProCys----- 994  
Db 2068 GCATCCGAACATCAACATGATCCTGAAAGGAACAGTACCTGTGTATAGGATGTGATACAG 2127  
QY 995 LeuPheHisArgSerProSerTyrIleValCysAsnThrThrSerSerAspGluVal 1014  
Db 2128 GTTAGCCATGTGCTAAATGACACCCACATGAATTCCTCTCCATCAACCGGAAA--- 2184  
QY 1015 LeuGluMetLys---ValSerValGlnValAspArgAlaLysIleHisGlnAspLeuVal 1033  
Db 2185 ---GAAATGAAGATGTGTATCCAGTTGTATGGTGGGAACCTCTCTCTGTGGGATCC 2241  
QY 1034 PheGlnTyrValGluAspProThrIleValArgIleGluProGluTrpSerIleValSer 1053  
Db 2242 TTATCTCATCTCTCTGCCACATTTGTCCTTATATTTCTGCTGCTACCATCGATCAGT 2301  
QY 1054 GlyAsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsnProGlnIle 1073  
Db 2302 GGTGGTCAAAATATACCATGATGGGAGCAAAATTTTGATGATTAATTGACAAC-----TTA 2355  
QY 1074 ArgAlaLysHisGlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGlu 1093  
Db 2356 ATCATTTTCATGAATTAATAAGGAACATAAATGCTCTCTGAATATTTGCTGGCGACTTAC 2415  
QY 1094 MetThrCysGlnAlaProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGlu 1113  
Db 2416 TCGCGGTTTTTACCCCCAGTTTA----- 2439  
QY 1114 ArgProGluGluPheGlyPheIleLeuAspAsnValGlnSerLeuLeuIleLeuAsnLys 1133  
Db 2440 -----AAGAGTTCAAAAGTGGCCACCAATGCTACTCTGGAAGTGAGAGTACAAGAC 2490  
QY 1134 Thr-----AsnPheThrTyrTrpProAsnProValPheGluAlaPheGly 1148  
Db 2491 ACCTACTTGTGATTTGTGGAACCCCTGCAGTATCGGAGGAGCCCAAGTTCAGCGGTATCGG 2550  
QY 1149 ProSerGlyIleLeuGlnLeuLysProGlyThrProIleIleLysGlyLysAsnLeu 1168  
Db 2551 -----GTGGAATCCGAGGTGGACACAGAACTGGAAGTGAATAATTCAA----- 2592  
QY 1169 IleProProValAlaGlyGlyAsnValLysLeuAsnTyrThrValLeuValGlyGluLys 1188  
Db 2593 -----AAAGAAATGACAACCTTCAATATTTTCCAAAAA 2625  
QY 1189 ProCysThrValThrVal-----SerAspValGlnLeuLeuCysGluSerProAsn 1205  
Db 1189 ProCysThrValThrVal-----SerAspValGlnLeuLeuCysGluSerProAsn 1205

Db 2626 GACATTGAAATTTACTCTCTCCATGGGAAATATGGCAATTAATTCAGTGTTCGAAAT 2685  
Qy 1206 Leulle- 1207  
Db 2686 ATTACTAGAAATCAAGATCTTACCACCATCTTTGCCAAATTAAGGCATCAAGACTGCA 2745  
Qy 1208 -----GlyArgHisValMetAlaArgValGlyGlyMetGluTyrSer 1222  
Db 2746 AGCACCATTGCCACTCTTCAAGAAAGTTCCGGTCAAGCTGGGAACCTGGAG----- 2799  
Qy 1223 ProGlyMetValTyrIleAlaProAspSer-----ProLeuSerLeuProAlaIleValSer 1241  
Db 2800 -----CTCTACGCTCGAGCAGAGTCAGTCTTCCATTCACATGTGATTTCTCTGATGTG 2850  
Qy 1242 IleAlaValAlaGlyGlyLeuLeuIlePheIleValAlaValLeuLeuAlaTyrLys 1261  
Db 2851 CTCCTCTGTC-----TTGCTAGTGATGTGCAATTTTGGCGCGTGGGGGTGACCAGG 2901  
Qy 1262 ArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsnLeu 1281  
Db 2902 CACAATCGAAGGAG-----CTGAGTCGCAACAGAGTCAA--CAACTAGAAATTCGTCG 2952  
Qy 1282 GluSerArgValAlaLeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIleHis 1301  
Db 2953 GAAACGAGCTCCGGAAGAGATAGCTGACGGCTTTGTGAGCTGCAGATGGATAAATTCG 3012  
Qy 1302 GluLeuThrSerAspLeuAspGlyAlaGly---IleProPheLeuAspTyrArgThrTyr 1320  
Db 3013 GATGTG-----GTTGATAGTTTGGAACTGTTCCTCTTCTTGACTACAAACATTTT 3063  
Qy 1321 ThrMetArgValLeuPhePro-----GlyIleGluAspHisProValLeuArgAspLeu 1338  
Db 3064 GCTCTGAGAACCTTCTTCCTCGAGTCAGTCAGTGGCTTCCACCACATCTTCACCTGAAGATATG 3123  
Qy 1339 GluValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAla----- 1356  
Db 3124 -----CATAACAGACAGCCCAACGACGAAGTGAAGTCTCACAGCTTGGAT 3171  
Qy 1357 GlnLeuLeuAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArg 1376  
Db 3172 GCCCTAATCTGTAATAAAGCTTCTTGTTACTGTCTACCCACACCCCTTGAAAGCAGAGA 3231  
Qy 1377 SerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSer 1396  
Db 3232 AACTTTTCTGTGAAGCAGAGGTCTGTGTTGCTCTCTTCTTACCAATTCACCTGCAAAACC 3291  
Qy 1397 LysLeuGluTyrAlaThrAspValLeuLysGlnLeuAlaAspLeuLeuAspLysAsn 1416  
Db 3292 AAGCTGGTCTACTGACACGATCTTAGAGGTGCTGACAGGAGCTGTATGGAAACAGTGT 3351  
Qy 1417 LeuGluSerLysAsnHisProLysLeuLeuArgThrGluSerValAlaGluLys 1436  
Db 3352 -----AGTAACATGCAGCCGAAACTCATGCTGAGACGCGAGTCCGTCGCGAATA 3405  
Qy 1437 MetLeuThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGlu 1456  
Db 3406 CTCCTCACAACTGGATGTCCTGCTCCCTTCTGATTTCTCCGGGAGAGCTGCGGAGAG 3465  
Qy 1457 ProLeuPheSerLeuPheCysAlaIleLysGlnGlnMetGluLysGlyProIleAspAla 1476  
Db 3466 CCCCTCTATTGCTGGTGAGACTCTGAACGAGAAATAACAAAGGTCCTCCGTCGGATGTA 3525  
Qy 1477 IleThrGlyGluAlaArgTyrSerLeuSerGluAspLysLeuLeuArgGlnIleAsp 1496  
Db 3526 ATCACTTGCAAGCCCTGTACACACTTAATGAAGACTGGCTGTGTGGCAGTTCGGAA 3585  
Qy 1497 TyrLysThrLeuValLeuSerCysVal-----SerProAspAsnAlaAsnSerPro 1513  
Db 3586 TTCAGTACTGTGGCATTAACCGTCTTTGAAAAAATCCCGGAAACAGAGTGGCAGAT 3645  
Qy 1514 -----GluValProValLysIleLeuAsnCysAspThrIleThrGlnValLysGlu 1530  
Db 3646 GTCTGTCGGAATATTTCAGTCAATGTCTCGACTGTGCACACCATTCGCGCAACCAAGAA 3705

Qy 1531 LysIleLeuAspAlaIlePheLysAsnValProCysSerHisArgProLysAlaAlaAsp 1550  
Db 3706 AGATTTTCCACGATCTTAAACCAAAATGGCTCTCTTATGAGCTTCAGCTTAATGAA 3765  
Qy 1551 MetAspLeuGluTrpArgGlnGlySerGlyAlaArgMetIleLeuGlnAspGluAspIle 1570  
Db 3766 ATTGGTCTTTCAGCTTCAAAATGGGCACACGACAGAAAGAACTTCTGGACATCGACAGTTC 3825  
Qy 1571 ThrThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHisTyrGlnValPro 1590  
Db 3826 TCCGCTGATCTTGAAGATGGAATCACCAAGCTAAACACCATTTGGCCACTATGACATATCA 3885  
Qy 1591 AspGlySerValValAlaLeuValSerLysGlnValThrAlaTyrAsnAlaValAsnAsn 1610  
Db 3886 AATGGATCCACTATAAAGTC----- 3906  
Qy 1611 SerThrValSerArgThrSerAlaSerLysTyrGluAsnMetIleArgTyrThrGlySer 1630  
Db 3907 -----TTTAAGAGATAGCAAAATTTTACTTCAGAT 3936  
Qy 1631 ProAspSerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyValLys 1650  
Db 3937 GTGAGTACTCGGATGACCACTGCCATTTGATTTTACAGATTCGGAAGCA----- 3987  
Qy 1651 MetTrpHisLeuValLysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGlySer 1670  
Db 3988 ---TTCCAAGATGTGCAAGGAAGAGACAT-----CGAGGAAG 4023  
Qy 1671 Lys-----MetValSerGluIleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeu 1688  
Db 4024 CACAAGTTCAAAGTAAAGAAATGTATCTGCAAAAGTGTCTGCAGCAAGGTGCAAT 4083  
Qy 1689 GluLysPheValAspAspLeuPheGluThrIlePheSerThrAlaHisArgGlySerAla 1708  
Db 4084 CATTCTGTGCTTGAAACACITTTTAGAGCATTTGGAGTTTACCCAAC-----AGCAGA 4137  
Qy 1709 LeuProLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGly 1728  
Db 4138 GCTCCATTTCTATAAAATACTTTTGTGACITTTTGGACGCCAGGCTGAAACAAAAA 4197  
Qy 1729 IleHisAspProHisValArgHisThrTrpLysSerAsnCysLeuProLeuArgPheTrp 1748  
Db 4198 ATCAGACTCTGACGTCGTACATATTTGGAAACAAACACCCCTCTCTTCGCTTCGG 4257  
Qy 1749 ValAsnMetIleLysAsnProGlnPheValPheAspIleHisLysAsnSerIleThrAsp 1768  
Db 4258 GTAACATCCTGGAAGAACCCCTCAGTCTTGTCTTGCATTTAAGAGACACACATATAGAC 4317  
Qy 1769 AlaCysLeuSerValValAlaGlnThrPheMetAspSerCysSerThrSerGluHisArg 1788  
Db 4318 GGTGTGTGTGTCAGTATGCGCCAGCATTCATGGATGATTTCTCTCACAGAGCAGCAA 4377  
Qy 1789 LeuGlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysAspIleProSerTyrLys 1808  
Db 4378 CTAGGGAAGGAGGACCAACCACTAATAAGCTTCTATGCCAGGATATCCCACTACAAA 4437  
Qy 1809 AsnTrpValGluArgTyrTyrSerAspIleGlyLysMetProAlaIleSerAspGlnAsp 1828  
Db 4438 GAAGAACTAAATCTTATTAACAAAGCAATCAGGATTTGCTCTCCATGTCATCCTCAGAA 4497  
Qy 1829 MetAsnAlaTyrLeuAlaGluInSerArgMetHisMetAsnGluPheAsnThrMetSer 1848  
Db 4498 ATGAAGAAATTTTAACTCAGAAATCTAAGAACATGAAATGAATTTAATGAAGAAGTG 4557  
Qy 1849 AlaLeuSerGluIlePheSerTyrValGlyLysTyrSerGluLeuIleLeuGlyProLeu 1868  
Db 4558 GCCTTGACAGAAATTTACAAATACATCGTAAATATTTTGTGACATCTAAATAACAT 4617  
Qy 1869 AspHisAsp-----AspGlnCysGlyLysGlnLysLeuAlaTyrLysLeu 1883  
Db 4618 GAAAGAACACGAGGCTGGAAGAGCTCAGAAACAACTCTTGCATGTAAAAAGTC 4671



QY	480	ProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGlu	499
Db	1359	-----TTTACGGACCANTCATGAG- ::: :::	1379
QY	500	ArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrClnSerCysGlyGluCys	519
Db	1379	----- ::: :::	1379
QY	520	LeuGlySerGlyAspProHisCysGlyTyrCysValLeuHisAsnThrCysThrArgLys	539
Db	1380	-----CACTGC-        :::	1409
QY	540	GluArg---CysGluArgSerLysGluProArgArg-        :::	556
Db	1410	TCATCAGCGGTGAAGCGCCGTGATGAATATCGAACAGAGTTTTACCACAGCTTTGCCAG ::: :::	1469
QY	557	GlnCysValArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeu	576
Db	1470	CgcgTT----- ::: :::	1505
QY	577	LeuVal-----LeuGluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThr	594
Db	1506	CTCTTAACACTATATCCACCTTCATTAAAGGAGACCTC- ::: :::	1547
QY	595	PheGluAspLeuSerGluMetAspGlyLeuValValcIyAsnGlnIleGlnCysTyrSer	614
Db	1548	ATAGCTAATCTTGGGACATCAGAGGTCGCTTCATGCAGGTTGTGTTCTTCGATCAGGA ::: :::	1607
QY	615	ProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAspHisHisValvalGln	634
Db	1608	CCATCAACCCCTCATGTGAATTTCTCCTG-----GACTCCCACATCCAGTG-- ::: :::	1652
QY	635	LeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerThrSerPheValPheTyr	654
Db	1652	----- ::: :::	1652
QY	655	AsnCysSerValHisAsnSerCysLeuSerCysValGluSerProTyrArgCysHisTrp	674
Db	1653	-----TCTCCA-        :::	1658
QY	675	CysLysTyrArgHisValCysThrHisAspProLysThrCysSerPheGlnGluGlyArg	694
Db	1658	----- ::: :::	1658
QY	695	VallLysLeuProGluAspCysProGlnLeuLeuArgValAspLysIleLeuValProVal	714
Db	1658	----- ::: :::	1658
QY	715	GluValIleLysProIleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGln	734
Db	1659	GAAGTGATTGGGACATACATTA-----AACCAA         :::	1688
QY	735	ArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArg	754
Db	1689	AATGGCTACACACTGGTTATCACT-----        :::	1712
QY	755	PheAsnSerSerSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIle	774
Db	1713	-----GGAGAAGATC ::: :::	1724
QY	775	AsnAsnLeuProValGluLeuThrValValTrpAsnGly-----HisPheAsn	790
Db	1725	ACGAAGATCCCCATTG-----AATGGCTGGCGCTGCAGACATATTC- ::: :::	1763
QY	791	IleAspAsnProAlaGlnAsnLysValHisLeuTyrLysCysGlyAlaMetArgGluSer	810
Db	1764	-----CAGTCC ::: :::	1769
QY	811	CysGlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTyrCysGlnGlyProGly	830
Db	1770	TGCAGTCAATGCCCTCTCTGCCCCACCCCTTTGTCAGTGTGCTGGTGCCAC-----GAC ::: :::	1823

QY	831	GlnCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuGluLeuSerGly	850
DB	1824	AAATCTGTGGATCGGAGGATGC-----CTCAGCGGG	1856
QY	851	AlaLysSerLys-----CysThrAsnProArgIleThrGluIleIleProValThrGly	868
DB	1857	ACATGGACTCAACAGATCTGCTG---CCTGCATCTACAAGTTTCCCAATAGTGCA	1913
QY	869	ProArgGluGlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPheArg	888
DB	1914	CCCTTGAAGGAGGACCAAGCTCACCATAATGTGCTGGGACTTTGGA-----TTTCGG	1967
QY	889	-----AspIleAlaSerHisValLysValAlaGlyValGluCysSerPro	903
DB	1968	AGGAATAATAAATTTGATTTAAAGAAACCTAGAGTCTCTCTTGGAAATGAGAGCTGCACC	2027
QY	904	LeuValAspGlyTyriIleProAlaGluGlnIleValCysGluMetGlyGluAlaLysPro	923
DB	2028	TTGACTTTTAAGTGAGACACCAATCAATCAATGAAATGCACAGTTGTT-----CCT	2078
QY	924	SerGlnHisAlaGlyPheValGluLleCysValAlaValCysArgProGluPheMetAla	943
DB	2079	GCCATGAATAAGCATTTTC---AATATGTCCATAATATTTCAAATGGCCAGGGACAACA	2135
QY	944	ArgSerSerGlnLeuTyriTyriPheMetThrLeuThrLeuSerAspLeuLysProSerArg	963
DB	2136	CAATACAGTACATCTCCTAT---GTGGATCCTGTAATAACAAGTATTTCGCCGAATAC	2192
QY	964	GlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySer	983
DB	2193	GGTCTATGGTGTGGGCACTTTACTTACTTAACTGGAAATTAACCTAAACAGTGGGAAT	2252
QY	984	AsnValValValMetPheGlyLysGlnProCysLeuPheHisArgSerProSerTyr	1003
DB	2253	TCTACACACATTTCAATTTGGTGGAAAACATGTACTTTAAAAGTGTGTCAACAGTATT	2312
QY	1004	IleValCysAsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnVal	1023
DB	2313	CTTGAATGTTATACCCAGGCCAAACCACTTTCAACTGAGTTGTGTGTTAAATTTGAAAT	2372
QY	1024	AspArgAlaLysIleHisGlnAspLeuValPheGlnTyrValGluAspProThrIleVal	1043
DB	2373	GACTTAGCCAAC---CGAGAGCAACGATCTTCAGTTACCGTGAAGATCCCATGTCTAT	2429
QY	1044	ArgIleGluProGluTrpSerIleVal-----	1053
DB	2430	GAATTCATCCAAACCAATCTTTTATAGTACTTGGTGGAAAGAACCTCTCAACATTGTC	2488
QY	1053	-----SerGlyAsnThrProIleAlaValTrpGlyThrHisLeu	1065
DB	2490	AGTTTTCTATTGTTTGCAGTGGTGGGACACATAACAGAGTGTGGGAAAACCTG	2549
QY	1066	AspLeuIleGlnAsnProGln-----IleArgAlaLysHisGlyLysGluHisIle	1088
DB	2550	AATTCAAGTTAGTCTCCGAGATGTCATAAATGTGCATGAAGCAGGAAGCACTTTACA	2605
QY	1084	AsnIleCysGluValLeuAsnAlaThrGlu---MetThrCysGlnAlaProAlaLeuAla	1107
DB	2610	GTGGCATGTCAACATCGCTCTAATTACAGAGATAATCTGTGTACCACCTCCCTCCCTG	2666
QY	1103	LeuGlyProAspHisGlnSerAspLeuThrGluArgProGluPheGlyPheIleLeu	1122
DB	2667	-----CAACACTGAATCTGCAACTCCCTCGAAACCAACAGCCTTT---TTCATGTTA	2711
QY	1123	AspAsnValGlnSerLeuLeuIleLeuAsnPheThrAsnPheThrTyrTyriProAsnPro	1144
DB	2718	GATGGGATCCTTTCCTCAAAATACCTTT-----GATCTCATTTATGTACATAATCCT	2765
QY	1143	ValPheGluAlaPheGlyProSerGlyIleLeuGluLeuLysProGlyThrProIleIle	1166
DB	2766	GTGTTTAAAGCCTTTGAAAAGCCAGTATGATCTCAATGGGGCAATCAAAATGTACTGGAA	2821
QY	1163	LeuLysGlyLysAsnLeuIleProProValAlaGlyGlyAsnValLysLeuAsnTyrThr	1181



```

QY 302 ArgSerGlyValGlu-----TyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGly 319
Db 1113 ACATCCACAAGAGGAAGTGTAAATATATCTACGGCTGCGTATCTGCAGCAAGCCTGGG 1172
QY 320 AlaValLeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPhe 339
Db 1173 GCCCAGCTTGCAGACAAATAGGACGACCGCTGAATGATGACATCTTTTCGGGGTGTTC 1232
QY 340 SerLysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIle 359
Db 1233 GCACAAAGCAGCCAGATCTGCCGAACCAAGATCGATCGATCGATGTCGATCCCT 1292
QY 360 LeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGluGly 379
Db 1293 ATCAATATGTCACGACTTCTTCAACAAG----- 1322
QY 380 ThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCysSerSerAlaLeuLeuThr 399
Db 1322 ----- 1322
QY 400 IleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyValSerAspMetVal 419
Db 1322 ----- 1322
QY 420 ArgGlyIleProValPheThrGluAspArgAspArgMetThrSerValIleAlaTyrVal 439
Db 1323 -----ATCGTC 1328
QY 440 TyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGlyLysLeuLysIleArg 459
Db 1329 ACACAAACAAATGTG----- 1343
QY 460 ValAspGlyProArgGlyAsnAlaLeuGlnTyrGluThrValGlnValValAspProGly 479
Db 1344 -----AGATGCTCCAGCAT----- 1358
QY 480 ProValLeuArgAspMetAlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGlu 499
Db 1359 -----TTTACGGACCCCAATCATGAG----- 1379
QY 500 ArgGlnLeuThrArgValProValGluSerCysGlyGlnTyrGlnSerCysGlyGluCys 519
Db 1379 ----- 1379
QY 520 LeuGlySerGlyAspProHisCysGlyTyrCysValLeuHisAsnThrCysThrArgLys 539
Db 1380 -----CACTGC-----TTTAATAGGACACTTCTGAGAAAT 1409
QY 540 GluArg--CysGluArgSerLysGluProArgArg-----PheAlaSerGluMetLys 556
Db 1410 TCATCAGGCTGTGAAGCGCGCTGATGAATATCGAACACAGAGTTTACCACAGCTTTCGAG 1469
QY 557 GlnCysValArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeu 576
Db 1470 CGCGTT-----GACTATTTCATGGTCAATTCAGCGAAGTC 1505
QY 577 LeuVal-----LeuGluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThr 594
Db 1506 CTCTTAACATCTATATCCACCTTCATTAAAGGAGACCTC-----ACC 1547
QY 595 PheGluAspLeuSerGluMetAspGlyLeuValValGlyAsnGlnIleGlnCysTyrSer 614
Db 1548 ATAGCTAATCTTGGGCATCAGAGGTCGCTTCATCAGAGTGTGGTTTCTCGATCAGGA 1607
QY 615 ProAlaAlaLysGluValProArgIleIleThrGluAsnGlyAspHisHisValValGln 634
Db 1608 CCATCAACCCCTCATGTGAATTTTCCCTG-----GACTCCCATCCAGTG----- 1652
QY 635 LeuGlnLeuLysSerLysGluThrGlyMetThrPheAlaSerThrSerPheValPheTyr 654
Db 1652 ----- 1652
QY 655 AsnCysSerValHisAsnSerCysLeuSerCysValGluSerProTyrArgCysHisTrp 674

```

---

```

Db 1653 ----- 1658
QY 675 CysLysTyrArgHisValCysThrHisAspProLysThrCysSerPheGlnGluGlyArg 694
Db 1658 ----- 1658
QY 695 ValLysLeuProGluAspCysProGlnLeuLeuArgValAspLysIleLeuValProVal 714
Db 1658 ----- 1658
QY 715 GluValIleLysProIleThrLeuLysAlaLysAsnLeuProGlnProGlnSerGlyGln 734
Db 1659 GAAGTGATGTGGAGCATACATTA-----AACCAA 1688
QY 735 ArgGlyTyrGluCysIleLeuAsnIleGlnGlySerGluGlnArgValProAlaLeuArg 754
Db 1689 AATGGCTACACACTGGTTATCAT----- 1712
QY 755 PheAsnSerSerSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIle 774
Db 1713 -----GGGAAGAAGATC 1724
QY 775 AsnAsnLeuProValGluLeuThrValTrpAsnGly-----HisPheAsn 790
Db 1725 ACGAAGATCCCATG-----AATGCTTGGGCTGCAGACATTTTC----- 1763
QY 791 IleAspAsnProAlaGlnAsnLysValHisLeuTyrLysCysGlyAlaMetArgGluSer 810
Db 1764 -----CAGTCC 1769
QY 811 CysGlyLeuCysLysLeuLysAlaAspProAspPheAlaCysGlyTrpCysGlnGlyProGly 830
Db 1770 TGCAGTCAATGCCYCTCTGCCCCACCCCTTTTTCAGTGTGGTGTGGCCAC-----GAC 1823
QY 831 GlnCysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTrpLeuLeuLeuSerGly 850
Db 1824 AATGTGTGCCATCGGAGGAATGC-----CTGACGCGG 1856
QY 851 AlaLysSerLys-----CysThrAsnProArgIleThrGluIleIleProValThrGly 868
Db 1857 ACATGGACTCAACAGATCTGTCTG--CCTGCATCTACAGGTTTCCCAAAATAGTGCA 1913
QY 869 ProArgGluGlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPheArg 888
Db 1914 CCCCTTGAAGGAGGACAAAGCTGACATATGTGCTGGGACTTTGGA-----TTTCGG 1967
QY 889 -----AspIleAlaSerHisValLysValAlaGlyValGluCysSerPro 903
Db 1968 AGGAATAATAATTTGATTTAAAGAAACTAGAGTTCTCCTTGGAAATGAGAGCTCCACC 2027
QY 904 LeuValAspGlyTyrIleProAlaGluGlnIleValCysGluMetGlyGluAlaLysPro 923
Db 2028 TTGACTTTAAGTGAGACGACGATGAATACATTTGAATGACATGACAGTTGGT-----CCT 2078
QY 924 SerGlnHisAlaGlyPheValGluIleCysValAlaValCysArgProGluPheMetAla 943
Db 2079 GCCATGAATAAGCATTTTC---AATATGTCATAATATTTTCAATAGCCACGGGACAAACA 2135
QY 944 ArgSerSerGlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysProSerArg 963
Db 2136 CAATACAGTACATTTCTCTAT--GTGGATCCTGTAATAACAAGTATTTTCGCCGAATAC 2192
QY 964 GlyProMetSerGlyGlyThrGlnValThrIleThrGlyThrAsnLeuAsnAlaGlySer 983
Db 2193 GGTCCTATGGTGTGGCAGCTTTACTTTTAACTGGAAATTTACTTAACAGTGGGAAT 2252
QY 984 AsnValValMetPheGlyLysGlnProCysLeuPheHisArgSerProSerTyr 1003
Db 2253 TCTAGACACATTTCAATTTGGTGGAAACATGTACTTTAAAGAGTGTGTCAACAGTATT 2312
QY 1004 IleValCysAsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnVal 1023

```

Db 2313 CTTGAATGTTATACCCCAACCAACCACTTCAACTGAGTTTCTGCTTTAAATGAAAT 2372  
QY 1024 AspArgAlaLysIleHisGlnAspLeuValPheGlnTyrValGluAspProThrIleVal 1043  
Db 2373 GACTTACCAAC--CGAGACAACAGCATCTTCAGTTACCGTGAAGATCCCATGCTAT 2429  
QY 1044 ArgIleGluProGluTyrSerIleVal----- 1052  
Db 2430 GAAATTCATCCCAACCAATCTTTATTAGTACTTGGTGAAGAACCTCTCAACATGTG 2489  
QY 1053 -----SerGlyAsnThrProIleAlaValTrpGlyThrHisLeu 1065  
Db 2490 AGTTTCTATTGTTGCTTCCAGATGTTGGGACACATAACAGGTGTTGGGAAACCTG 2549  
QY 1066 AspLeuIleGlnAsnProGln-----IleArgAlaLysHisGlyGlyLysGluHisIle 1083  
Db 2550 AATTCAGTTAGTCCCGAGATGTCATAATGTCATGAACGACGAAGAACTTTACA 2609  
QY 1084 AsnIleCysGluValLeuAsnAlaThrGlu---MetThrCysGlnAlaProAlaLeuAla 1102  
Db 2610 GTGGCATGTCACATCGCTCTAATTCAGAGATAATCTGTTGACCACTCCTTCCCTG--- 2666  
QY 1103 LeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPheGlyPheIleLeu 1122  
Db 2667 -----CAACAGCTGAATCGCACTCCCTCGGAAACCAACGCTTT---TTCATGTTA 2717  
QY 1123 AspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsnPheThrTyrTyrProAsnPro 1142  
Db 2718 GATGGGATCCTTTCCAATACITTT-----GATCTCATTTATGTACATATCCT 2765  
QY 1143 ValPheGluAlaPheGlyProSerGlyIleLeuGluLeuLysProGlyThrProIleIle 1162  
Db 2766 GTGTTTAAGCCVTTTGAAGAGCCAGTGATGATCTCAATGGCAATGAAATGTAAGTGA 2825  
QY 1163 LeuLysGlyLysAsnLeuIleProProValAlaGlyGlyAsnValLysLeuAsnTyrThr 1182  
Db 2826 ATTAAGGAAATGATATGACCTGACCTGAAGCAGTTAAGGTGAAGT----- 2870  
QY 1183 ValLeuValGlyGluLysProCys---ThrValThrValSerAspValGlnLeuLysCys 1201  
Db 2871 TTAAGAGTTGAATAAGAGCTGTGAGATATACACTTACATCTCAAGCCGTTTATGC 2930  
QY 1202 GluSerProAsn---LeuIleGlyArgHisLysValMetAlaArgValGlyGlyMetGlu 1220  
Db 2931 ACGGTCCCAATGACCTGCTGAATTAAGACGAGCTA-----AATATAGAG 2978  
QY 1221 TyrSerPro-----GlyMetValTyrIleAlaProAspSerPro 1233  
Db 2979 TGAAGCAAGCAATTTCTTCAACCGTCCCTGGAAAGTAATAGTTCACACGATAGAA 3038  
QY 1234 LeuSer-----LeuProAlaIleValSerIleAlaValAlaGlyGlyLeuIleIle 1251  
Db 3039 TTCACAGGATTCATGCTGCTGCTCAATATCAACAGCAGCTGTTATTACTACTGGG 3098  
QY 1252 PheIleValAlaValLeuIleAlaTyrLysArgLysSerArgGluSerAspLeuThrLeu 1271  
Db 3099 TTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3158  
QY 1272 LysArgLeuGluMetClnMetAspAsnLeuGluSerArgValAlaLeuGluCysLysGlu 1291  
Db 3159 TACGATGCAAGAGTACACACTCCTCATTTGGATGAGCTGTTGTAAGTCCCGAAGTGAAG 3218  
QY 1292 AlaPheAlaGluLeuGlnThrAspIleHisGluLeuThrSerAspLeuAspGlyAlaGly 1311  
Db 3219 CCAACTACAGAAATGGTTTCAAT----- 3242  
QY 1312 IleProPheLeuAspTyrThrThrTyrThrMetArgValLeuPheProGlyIleGluAsp 1331  
Db 3243 ---GAATCTGTAGACTACCGAGCTACTTTTCCAGAGATCAGTTTCTTAATTCATCTCAG 3299  
QY 1332 HisProValLeuArgAspLeuGluValPro 1341  
Db 3300 AACGGTTCATGCCGACAAAGTCAGTATCCT 3329

## RESULT 6

US-08-121-713D-61  
: Sequence 61, Application US/08121713D  
: Patent No. 5639856  
: GENERAL INFORMATION:  
: APPLICANT: Goodman, Corey S.  
: APPLICANT: Kolodkin, Alex L.  
: APPLICANT: Matthews, David  
: APPLICANT: Bentley, David R.  
: APPLICANT: O'Connor, Timothy  
: TITLE OF INVENTION: The Semaphorin Gene Family  
: NUMBER OF SEQUENCES: 100  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
: STREET: 268 Bush Street, Suite 3200  
: CITY: San Francisco  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/121,713D  
: FILING DATE: 13-SEP-1993  
: CLASSIFICATION: 514  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Osman, Richard A.  
: REGISTRATION NUMBER: 36,627  
: REFERENCE/DOCKET NUMBER: B94-002-1  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 343-4341  
: TELEFAX: (415) 343-4342  
: TELEX:  
: INFORMATION FOR SEQ ID NO: 61:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 2670 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 268..2439  
: US-08-121-713D-61

Alignment Scores:  
Pred. No.: 1,47e-13 Length: 2670  
Score: 227.00 Matches: 147  
Percent Similarity: 37.22% Conservat: 121  
Best Local Similarity: 20.42% Mismatches: 259  
Query Match: 2.27% Indels: 194  
DB: 1 Gaps: 34

US-09-964-956-13 (1-1896) x US-08-121-713D-61 (1-2670)

QY 53 HisLeuValValAspGluArgThrGlyHisIleTyrLeuGlyAlaValAsnArgIleTyr 72  
Db 481 CATATGAACCAAGATCAGATACG-----CTCTATGTGGGAGCCATGATCGCGTATTC 534  
QY 73 LysLeuSerSerAspLeuLysValLeuValThrHisGluThrGlyProAspClnAspAsn 92  
Db 535 CGTGTG-----AACCTGCAGAAATATCTCTCAATTCATTCGGGATCGCATCAAC 588  
QY 93 ProLysCysTyrProProArgIleValGlnThrCysAsnGluProLeuThrThrAsn 112  
Db 589 -----TTGGAGCCCAACCGGATGATGTGTTAGTCCGCTC 624  
QY 113 AsnValAsnLysMetLeuLeuIleAspTyrLysGluAsnArgLeuIleAlaCysGlySer 132  
Db 113 AsnValAsnLysMetLeuLeuIleAspTyrLysGluAsnArgLeuIleAlaCysGlySer 132

Db	625	TCCAAGGCAAAAGTCTGAGTCTTCGACTGCAAG---AACCATGTGGCTGTCTATCCAGTCA	681
Qy	133	LeuTyrGlnGlyIleCysLysLeuLeuArgLeuGluAspLeuPheLysLeuGlyGluPro	152
Db	682	ATGGACCAAGGG---GATAGGCTCTATGTATCGGCACCAAC	720
Qy	153	TyrHisLysLysGluHisTyr-----LeuSerGlyValAsnGluSerGly	167
Db	721	GCCCAACAATCCCAAGGATTATGTATATGCAATCTAACCCACCTCGCGCGCTCGAA	780
Qy	168	SerValPheGlyValIleValSer-----TyrSerAsnLeuAspLys	182
Db	781	TATGTGATTCGGTGGCTTGGGCATTGCCAAGTCCCTACGATCCCTCGACAAC---	837
Qy	183	LeuPheIleAlaThrAlaVal-----AspGlyLysProGluTyrPheProThrIle	199
Db	838	-----TCAACTGGCATTTATGTGGAGAATGGCAATCGGGTGTCTGCCGGTTG	888
Qy	200	SerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHis	219
Db	889	TACTCGGCACCAATCGGAGTTCCACCAAGCGGAT-----ACGTTATTTTCGC	939
Qy	220	AspGluPheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAsp	239
Db	940	ACTGATCTGTATATACTTCGGCTAAACGTTTGGAAATATAAATTCACAGAGACTCTGAA	999
Qy	240	PheAspIle-----TyrTyrValTyrGlyPheSerSerGlyAsnPhe	253
Db	1000	TACGACTCCAAGTGGTGGACAAACCAACTTTTCGGCTCCTTTGATATTGGGAGTAC	1059
Qy	254	ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysG1	273
Db	1060	GTGTATTTCTTTTCGGTGAACCG-----CCGTGGAATACATCACTCGCGC	1107
Qy	273	uGlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAla-----	289
Db	1108	AAGGCTGTC-TATTTCGGCATCGCAGCGGTGTGCAAGAAGTGTGGGTGGAAGAATCT	1166
Qy	290	-----PheAsnSerTyrValGluValProIleGlyCysGluArgSerGly--	304
Db	1167	GCTGGCCACAACTTGGCCACCTACTTGAGAGCCAGACTCACTGCACATCTCGCGCA	1226
Qy	305	-----ValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaVal	322
Db	1227	ATTTCCGTTCTATTCAACGAGATCCAATCGTCTAC-----	1263
Qy	322	uGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerLysG1	342
Db	1264	-----CAGCTGCCCTCCGATAAGAGTCAATCTTCGCCATTCACG-----	1305
Qy	342	yGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysG1	362
Db	1306	-----ACGAGCACTAATGGCTGATTTGGATCTGCCGTATGCAGTTTCCACATTAACA	1358
Qy	362	nile-----AsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyG1	378
Db	1359	GATTTCAGGCTGCTTCAATGGCAAAATCAAGGAGCAATCTTCATCG-----	1404
Qy	378	uGlyThrLeuAspLeuAlaTrpLeuLysValLysAsp-----IleProCysSerSerAl	396
Db	1405	-----AATTCGGATGCTGCCGGTGTAACTCCCGGTGGCGGACCAACCGGCC	1454
Qy	396	aLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyVal	416
Db	1455	GGGTACATGTGTCAACGATACA-----TCAAAACCTGCCCGATACCGTACT	1499
Qy	416	rAspMetValArgGlyIlePro-----	423
Db	1500	GAATTTTCATCATGATCCCATCCATCTTATGGACAAAGCCGTAATTCAGACACAAATCC	1559
Qy	424	-----ValPheThrGlu-----AspArgAspArgMetTh	433
Db	1560	AGTCTATTATAAAGGATTGTCTCTTCAACCAAGCTCGTGTGACAAAATTCG-----	1614

Qy	433	rSerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerG1	453
Db	1615	-----ATTGCATCTCAACCAAGGATACATTGTGTACTATGTGGCACCACCAATCTGG	1667
Qy	453	yLysLeuLysLysIleArgValAspGlyProArgGlyAsnAlaLeu---GlnTyrGluTh	472
Db	1668	TCGCATTTACAAAATCGTGCAGTACTACCGTAACGGAGAGTGCCTGTCTCAAGCTTCTGA	1727
Qy	472	rValGlnValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisG1	492
Db	1728	TATCTTCAGGTGGCTCCAAACAGGCGCATCAAGTATGGAAATCAGCAGACGCTAA	1787
Qy	492	uGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCysGly--	511
Db	1788	GAGCCTCTACATTGGCACCAGCATCATCGATCAAGCAAAATCGACCTGGCCATGTGCAATCG	1847
Qy	512	-GlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTrpCysVa	531
Db	1848	CCGTTTACGACAACTGCTTCGCTGGGTC-----CGTGATCCCTACTGGCGTGGGATAA	1901
Qy	531	lLeuHisAsnThrCysThrArgLysGluArg-----	541
Db	1902	GGAGCCCAATACGTCGCGACCTACGAGCTGGATTCTACAGGATGTGCCAATGAAAC	1961
Qy	542	-----CysGluArgSerLysGluProArgArgPheAlaSerGluMetLysGlnCy	558
Db	1962	GAGTGACATTTCCGATTCGAGTGTCTGAAAAAGAGATTGTGGTGCACCTATGCCAGAG	2021
Qy	558	sValArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuLeuVa	578
Db	2022	TGTACATCTGGCGCTGTTTC-----	2040
Qy	578	lLeuGluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLe	598
Db	2041	-----GTCANAATACCGAAGTGTGGAAGATGAGCAAGTGCCTGGTATCATCA	2090
Qy	598	uSerGluMetAspGlyLeuValGlyAsnGlnIleGlnCysTyrSerProAlaAlaLy	618
Db	2091	CTCCAAGGACAAAGGACGCTAC-----GAGATTCTGT---TACTCGCCGACCAAAA--	2136
Qy	618	sGluValProArgIleThrGluAsnGlyAspHisHisValValGlnLeuGlnLeuLy	638
Db	2137	-----TACATTGAGACCCAGCCGACCTGGCTGTGTGTTTCCGTGAA	2180
Qy	638	sSerLysGluThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerVa	658
Db	2181	CGAAGCCGATGGTGGTCTCG-----TACGATTGCCCATTT	2213
Qy	658	lHisAsnSerCysLeuSerCysValGluSerProTyrArgCysHisTrpCysLysTyrAr	678
Db	2214	GGCGGCTCGCTTTTG-----TGCACGTACAA	2240
Qy	678	g-----HisValCysThrHisAspProLysThrCysSerPheGlnGlu	692
Db	2241	CATTACAGTGGATGCCACAGATGCATCTCCGCCGAACAAGAGTAATGACTATCAGAAA	2298

RESULT 7  
US-08-835-268-61  
Sequence 61, Application US/08835268  
Patent No. 5807826  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey S.  
APPLICANT: Kolodkin, Alex L.  
APPLICANT: Mathies, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family  
NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco

STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/835,268  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4341  
TELEFAX: (415) 343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2670 base pairs  
TYPE: nucleic acid  
STRADEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 268..2439  
US-08-835-268-61

Alignment Scores:  
Pred. No.: 1,47e-13 Length: 2670  
Score: 227.00 Matches: 147  
Percent Similarity: 37.22% Conservative: 121  
Best Local Similarity: 20.42% Mismatches: 259  
Query Match: 2.27% Indels: 194  
DB: 1 Gaps: 34

US-09-964-956-13 (1-1896) x US-08-835-268-61 (1-2670)

QY	183	LeuPheileAlaThrAlaVal-----AspGlyLysProGluTyrPheProThrile	199
DB	838	-----TCAACTGCGATTATGTGAGAAATGCGGTTGCTGCCCGGTTG	888
QY	200	SerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHis	219
DB	889	TACTCCGGCCCAATGCGGAGTTCCACCAAGCGGAT-----ACGGTTATTTTCGCG	939
QY	220	AspGluPheValAlaLysMetIleLysIleProSerAspThrPheIleIleProAsp	239
DB	940	ACTGATCTGTATAATACTTCGGCTAAACGTTTGAATAATAATTCACAGAGACTCTGAAA	999
QY	240	PheAspIle-----TyrTyrValTyrGlyPheSerSerGlyAsnPhe	253
DB	1000	TACGACTCCAAGTGGTGGACAAACAACTTGTTCGGCTCCTTTGATATTTGGGAGTAC	1059
QY	254	ValTyrPhe-LeuThrLeuGlnProGluMetValSerProProGlySerThrThrLysG1	273
DB	1060	GTGTATTTCTTTTTCGGTGAACCG-----CCGTGGAATACATCAACTGCGGC	1107
QY	273	uGlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAla-----	289
DB	1108	AGGCTGTC-TATTCCGCCATCGCAGCGGTGTCAAGAAGGATGTGGTGGAAAGAACT	1166
QY	290	-----PheAsnSerTyrValGluValProIleGlyCysGluArgSerGly--	304
DB	1167	GCTGGCCCAACACTGGCCACCTACCTGAAGCCAGACTCAACTGCAGCATCTCCGCGCA	1226
QY	305	-----ValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaValLe	322
DB	1227	ATTTCCGTTCTATTTCACAGAGATCCATCGGTCTAC-----	1263
QY	322	uGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerLysG1	342
DB	1264	-----CAGCTGCCCTCCGATAAGAGTCGATTCTTCGCCCATTCACG-----	1305
QY	342	yGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysG1	362
DB	1306	-----ACGAGCACTAATGCCCTGATTGGGATGCGCGTATGCACTTTCCACATTAACGA	1358
QY	362	nIle-----AsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyG1	378
DB	1359	GATTACGGCTGCCTTCAATGGCAATTCACAGGACCAATTTCTCATCG-----	1404
QY	378	uGlyThrLeuAspLeuAlaTyrLeuLysValLysAsp-----IleProCysSerSerAl	396
DB	1405	-----AATTCCGATGCGTCCGCGTCTTAATCTCCCGGGTGGCGGAACACAGGCC	1454
QY	396	aLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyValSe	416
DB	1455	GGGTACATGTGTCAACGATACA-----TCAAACCTGCCGATACCGTACT	1499
QY	416	rAspMetValArgGlyIlePro-----	423
DB	1500	GAATTTTCATCAGATCCCATCCACTTATGGCAAAAGCCGTAATCACAGACACACAATCC	1559
QY	424	-----ValPheThrGlu-----AspArgAspArgMetTh	433
DB	1560	AGTCTATTATAAAGGATTTGGTCTTCACCAAGCTCGCTGTCGACAAAATTCGC-----	1614
QY	433	rSerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerG1	453
DB	1615	-----ATTCACATCTCTCAACAGGAATACATGTGTACTATGTGGGCACCAATCTGGG	1667
QY	453	yLysLeuLysLysIleArgValAspGlyProArgGlyAsnAlaLeu-----GlnTyrGluTh	472
DB	1658	TCGCATTTACAAAATCTGTCAGTACTACCGTAACGAGAGTCTGCTGTCACAGCTTCGGA	1727
QY	472	rValGlnValValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisG1	492
DB	1728	TATCTTCAGGTGCTCCAAACGAGGCATCCCAAGTATGTAAGAAATCAGCCAGACACGTAA	1787
QY	492	uGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCysGly--	511



Db 1000 TAGGACTCAAGTGGTGGACAAACAACTTTGTCGGCTCCTTTGATATTTGGGGAGTAC 1059  
QY 254 ValTyrPhe-LeuThrLeuGlnProGluMetValSerProProGlySerThrThrLysG1 273  
Db 1060 GTGTATTTCTTTTTCGTTGAAACCG-----CGTGAATACATCACTCGGC 1107  
QY 273 uGlnValTyrThrSerLysLeuValArgLeuGlyCysLysGluAspThrAla----- 289  
Db 1108 AAGGTGTC-TATTCGCGCATCGACGGGTGCAAGAAGATGGTGGTGAAGAATCT 1166  
QY 290 -----PheAsnSerTyrValGluValProIleGlyCysGluArgSerGly-- 304  
Db 1167 GCTGGCCACACACTGGCCACCTACTGAAGCCACACTCACTGAGCATCTCCGCGCA 1226  
QY 305 -----ValGluTyrArgLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaValLe 322  
Db 1227 ATTTCGGTTCTATTTCACGAGATCCAAATCGGPTAC----- 1263  
QY 322 uGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerLysG1 342  
Db 1264 -----CAGCTCCCTCCGATAGAGTCTGCTTCGCCACATTCACG----- 1305  
QY 342 yGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysG1 362  
Db 1306 -----ACGACACTAATGGCTGATTCGATCTGCCGATGCTTCACATTAACGA 1358  
QY 362 ntle-----AsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyG1 378  
Db 1359 GATTCAGGCTGCTTCAATGGCAATTCAGGAGCAATCTTCATCG----- 1404  
QY 378 uGlyThrLeuAspLeuAlaTrpLeuLysValLysAsp-----IleProCysSerSerAl 396  
Db 1405 -----AATTCGCGATGCTGCGGTCTTAATCTCCCGGTGCCGACACAGGCC 1454  
QY 396 aLeuLeuThrIleAspAsnPhcCysGlyLeuAspMetAsnAlaProLeuGlyValSe 416  
Db 1455 GGTACATGTGCAACCATACA-----TCAAACTGCCCGATACCGTACT 1499  
QY 416 rAspMetValArgGlyIlePro-----ValPheThrGlu-----AspArgAspArgMeth 433  
Db 1500 GAATTTTCATCAGATCCATCCACTTATGGACAAAGCCGTAATACGAGCACAAATCC 1559  
QY 424 -----ValPheThrGlu-----AspArgAspArgMeth 433  
Db 1560 AGTCATTATAAAGGATTTGCTCTCCAAAGCTCGTCTGTGACAAATTCG----- 1614  
QY 433 rSerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerG1 453  
Db 1615 -----ATTGACATCTCAACGAGTAATACATGTGTACTATGTGGCCCAATCTGG 1667  
QY 453 yLysLeuLysLysIleArgValAspGlyProArgGlyAsnAlaLeu---GlnTyrGluTh 472  
Db 1668 TCGCATTTACAAAATCGTCAGTACTACCGTAACGAGAGTCTGTCCTCAAGCTTCTGGA 1727  
QY 472 rValGlnValValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisG1 492  
Db 1728 TATCTTCGAGGTGGCTCCAAACGAGCCCATCCAAAGTGGAATACGACGACAGCATAA 1787  
QY 492 uGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCysGly-- 511  
Db 1788 GAGCCTCTACATGGCACCAGCATCGCATCAAGCAATCGACCTGGCCATGTGCAATCG 1847  
QY 512 -GlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTrpCysVa 531  
Db 1848 CCGTTACGACAACTGCTTCCGCTGCCTC-----CGTGATCCCTACTGCGCTGGGATAA 1901  
QY 531 lleuHisAsnThrCysThrArgLysGluArg----- 541  
Db 1902 GGAGGCCAATACGTGCGCCAGCTACGAGTGGATTTTACTGCAGAGTGTGGCCAAATGAAAC 1961  
QY 542 -----CysGluArgSerLysGluProArgArgPheAlaSerGluMetLysGlnCy 558  
Db 1962 GAGTGACATTTGCGATTCAGTGTGCTGAAAAGAGATGTGCTGACCTATGCGCCAGAG 2021

QY 558 sValArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuVa 578  
Db 2022 TGTACATCTGGCTGTTTC----- 2040  
QY 578 lLeuGluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLe 598  
Db 2041 -----GTCAAAATACCCGAAGTCTGAAGAAATGAGCAAGTGCCTGATCATCA 2090  
QY 598 uSerGluMetAspGlyLeuValValGlyAsnGlnIleGlnCysTyrSerProAlaAlaL 618  
Db 2091 CTCAAGACGACGAGCGCTAC-----GAGATTGCT---TACTCGCCGACCAAAA-- 2136  
QY 618 sGluValProArgIleThrGluAsnGlyAspHisValValGlnLeuGlnLeuL 638  
Db 2137 -----TACATTGAGACCCAGCAAGCTGCGCTGTTGTTGTTCCGTGAA 2180  
QY 638 sSerLysGluThrGlyMetThrPheAlaSerThrSerPheValPheTyrAsnCysSerVa 658  
Db 2181 CGAAGCCGATGGTGGTCCG-----TACGATTGCGCATTT 2213  
QY 658 lHisAsnSerCysLeuSerCysValGluSerProTyrArgCysHisTyrCysLysTyAr 678  
Db 2214 GGGCGGCTCGCTTTG-----TGCAGCTACAA 2240  
QY 678 g-----HisValCysThrHisAspProLysThrCysSerPheGlnGlu 692  
Db 2241 CATTACAGTGGATGCCACACAGATCCCTCCCGCAACAGAGTAATGACTATCAGAAA 2298

## RESULT 9

US-08-833-391-61  
; Sequence 61, Application US/08833391  
; Patent No. 6013781  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S.  
; APPLICANT: Kolodkin, Alex L.  
; APPLICANT: Matthes, David  
; APPLICANT: Bentley, David R.  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/833.391  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/121,713  
; FILING DATE: 13-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B94-002-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)343-4341  
; TELEFAX: (415) 343-4342  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2670 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double

; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 268..2439  
 ; US-08-833-391-61

## Alignment Scores:

Pred. No.: 1,47e-13 Length: 2670  
 Score: 227.00 Matches: 147  
 Percent Similarity: 37.22% Conservative: 121  
 Best Local Similarity: 20.42% Mismatches: 259  
 Query Match: 2.27% Indels: 194  
 DB: 3 Gaps: 34

US-09-964-956-13 (1-1896) x US-08-833-391-61 (1-2670)

QY 53 HisLeuValValAspGluArgThrGlyHisIleTyrLeuGlyAlaValAsnArgIleTyr 72  
 DDb 481 CATATGAACGAAGATCGAGATACG-----CTCTATGTGGGAGCCATGGATCGGTATTC 534  
 QY 73 LysLeuSerSerAspLeuLysValLeuValThrHisGluThrGlyProAspGluAspAsn 92  
 DDb 535 CGTGTG-----AACCTGCGAATATCTCTCATCCAAATTTGTAATCGGATCGCATCAAC 588  
 QY 93 ProLysCysTyrProProArgIleValGlnThrCysAsnGluProLeuThrThrAsn 112  
 DDb 589 -----TTGGAGCCACACGGGATGATGGTTAGTGGCTC 624  
 QY 113 AsnValAsnLysMetLeuLeuIleAspTyrLysGluAsnArgLeuLeuIleAlaCysGlySer 132  
 DDb 625 TCCAAAGCAAAAGTCAGATCTTCGACTGCAAG---AACCATGTGCGGTGCATCCAGTCA 681  
 QY 133 LeuTyrGlnGlyIleCysLysLeuLeuArgLeuGluAspLeuPheLysLeuGlyCluPro 152  
 DDb 682 ATGACACAGGG-----GATAGGCTCTATGTATGCGGCACCAAC 720  
 QY 153 TyrHisLysLysGluHisTyr-----LeuSerGlyValAsnGluSerGly 167  
 DDb 721 GCCACAAATCCAGGATTATGTATCTATCGGAATCTTAACCCACTCCCGCTCGGAA 780  
 QY 168 SerValPheGlyValIleValSer-----TyrSerAsnLeuAspLys 182  
 DDb 781 TATGTGATGGCGTGGCTGGGATTCGCCAAGTCCCTACGATCCCTCGACAAAC--- 837  
 QY 183 LeuPheIleAlaThrAlaVal-----AspGlyLysProGluTyrPheProThrIle 199  
 DDb 838 -----TCAACTGCGATTATGTGGAGAAATGCAATCCGGGTGCTGCCCGTTTG 888  
 QY 200 SerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHis 219  
 DDb 889 TACTCCGCGACCAATGCGGAGTTACCAAGGCGGAT-----ACGGTTATTTTCGCG 939  
 QY 220 AspGluPheValAlaSerMetIleLysIleProSerAspThrPheThrIleIleProAsp 239  
 DDb 940 ACTGATCTGTATAACTTACTTCGGCTAAACGTTTGGAAATATAAATTAACAGGACTCTGAAA 999  
 QY 240 PheAspIle-----TyrTyrValTyrGlyPheSerSerGlyAsnPhe 253  
 DDb 1000 TACGACTCCAAGTGGTTCGACAAACCAACTTTGTTCGGCTCCTTTGATATTGGGGAGTAC 1059  
 QY 254 ValTyrPhe-LeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGly 273  
 DDb 1060 GTGATATTCTTTTCCGTGAAACCG-----CCGTGGAATACATCAACTCGCGC 1107  
 QY 273 uGlnValTyrThrSerLysLeuValArgLeuLysGlyValAspThrAla----- 289  
 DDb 1108 AAGCGTGTCT-TATTCGCGCATCGCAGCGGTGTCGAAGAGGATGCGGTGGGAAGAAATCT 1166  
 QY 290 -----PheAsnSerTyrValGluValProIleGlyCysGluArgSerGly-- 304  
 DDb 1167 GCTGCCCCACAACTGGGCCACCTACCTGAAGCCGAGACTCAACTGAGCATCTCCGCGCA 1226

QY 305 -----ValGluTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaValLe 322  
 DDb 1227 ATTTCCGTTCTATTTCACGAGATCAATCGTCTAC----- 1263  
 QY 322 uGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerLysGly 342  
 DDb 1264 -----CAGCTGCGCTCCGATAAGAGTCTCTTCGGCCACATTCACG----- 1305  
 QY 342 yGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLysGly 362  
 DDb 1306 -----ACGAGCACTAATGCGCTGATTCGGATCTCCGCTATCCAGTTCCTCCATTAACGA 1358  
 QY 362 nile-----AsnAspArgIleLysGluArgLeuGlnSerCysTyrArgGlyGly 378  
 DDb 1359 GATTCAAGCTGCTTCAATGCAAAATCAAGAGCAATCTTCATCG----- 1404  
 QY 378 uGlyThrLeuAspLeuAlaTrpLeuLysValLysAsp-----IleProCysSerSerAl 396  
 DDb 1405 -----AATTCGCGATGCTCCGCTGCTTAACCTCCGCTGCGGAGCCGAAACCGGCC 1454  
 QY 396 aLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeuGlyValse 416  
 DDb 1455 GGGTACATGTGTCAACGATACA-----TCAACCTCGCCGATACCGCTACT 1499  
 QY 416 rAspMetValArgGlyIlePro----- 423  
 DDb 1500 GAATTTTCATCATGCCATCCATCCATTTATGACAAAGCCGTAAATCACAGCACAAATCC 1559  
 QY 424 -----ValPheThrGlu-----AspArgAspArgMetTh 433  
 DDb 1560 ACTCTATTATAAAGGGATTTGGTCTTCCAAAGCTCGTGTGACAAATTCGC----- 1614  
 QY 433 rSerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
 DDb 1615 -----ATTGACATCTCAACACGAGATACATTGTGTACTATGTGGCCACCAATCTGGG 1667  
 QY 453 yLysLeuLysLysIleArgValAspGlyProArgGlyAsnAlaLeu---GlnTyrGluTh 472  
 DDb 1668 TCGCATTTTCAAAATTCGTGCAGTACTACCGTACCGGAGAGTCTGCTTCCAACTTCTGGA 1727  
 QY 472 rValGlnValValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisGly 492  
 DDb 1728 TATCTTCGAGGTGGCTCCAAACGAGGCCATCCCAAGTATGAGAAATCAGCCAGACACGTAA 1787  
 QY 492 uGlnLeuTyrIleMetSerGluArgGlnLeuThrArgValProValGluSerCysGly-- 511  
 DDb 1788 GAGCCTCTACATTGGCACCAGCATCGCATCAAGCAAAATCGACTGGCCATGTGCAATCG 1847  
 QY 512 -GlnTyrGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTrpCysVa 531  
 DDb 1848 CGGTTACGACACTGCTTCCGCTGCGTC-----CGTGATCCCTACTGCGCTGGGATAA 1901  
 QY 531 lLeuHisAsnThrCysThrArgLysGluArg----- 541  
 DDb 1902 GGAGGCCAATACGTCGCGACCGTACGAGCTGGATTACTCAGGATGTGCCAATGAAC 1961  
 QY 542 -----CysGluArgSerLysGluProArgArgPheAlaSerGluMetLysGlnCy 558  
 DDb 1962 GAGTGACATTTCGATTCGAGTGTGTGAAAAAGAGATTGTGTGACCTATGCCCCAGAG 2021  
 QY 558 sValArgLeuThrValHisProAsnAsnIleSerValSerGlnTyrAsnValLeuLeuVa 578  
 DDb 2022 TGTACATCTGGCTGTTTC----- 2040  
 QY 578 lLeuGluThrTyrAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLe 598  
 DDb 2041 -----GTCAAAATACCCGAAGTCTGAGAATGAGCAATGAGCAATGCTGTATCAPCA 2090  
 QY 598 uSerGluMetAspGlyLeuValValGlyAsnGlnIleGlnCysTyrSerProAlaAlaLy 618  
 DDb 2091 CTCGAAGGACAAAGGAGCGTAC-----GAGATTCTGT---TACTCGCGCAGCAAAA-- 2136  
 QY 618 sGluValProArgIleIleThrGluAsnGlyAspHisValValGlnLeuGlnLeuLy 638



Db 1359 GATTAGGCTGCCTTCAATGGCAAAATTCAGGAGCAATCTTCATCG----- 1404

Qy 378 uGlyThrLeuAspLeuAlaTrpLeuLysAsp-----IleProCysSerSerAl 396

Db 1405 -----AATCCGATGGTCCGGTCTTAACCTCCGGTGGCGGAACACGCGCC 1454

Qy 396 aLeuLeuThrIleAspAspAspPheCysGlyLeuAspMetAsnAlaProLeuGlyValse 416

Db 1455 GGGTACATGTGTCAACGATACA-----TCAAACCTGCCCGGATACCGTACT 1499

Qy 416 rAspMetValArgGlyIlePro----- 423

Db 1500 GAATTTTCATCAGATCCCATCCACATTTATGGCAAAACCGCTAATCAGGAGCACACAATCC 1559

Qy 424 -----ValPheThrGlu-----AspArgAspArgMetTh 433

Db 1560 AGTCTATTATAAAGGGATTTGGTCTTCACCAAGCTCGTGTGACAAAATTCGC----- 1614

Qy 433 rSerValIleAlaTrpValTrpLysAsnHisSerLeuAlaPheValGlyThrLysSerGl 453

Db 1615 -----ATTGACATCTCTCAACAGGAATATGTGTACTATGGGACCAATCTGGG 1667

Qy 453 yLysLeuLysLysIleArgValAspGlyProArgGlyAsnAlaLeu---GlnTrpGluTh 472

Db 1668 TCGCATTTACAAAATCGTCGACTACTACCGTAACGAGAGTGGCTGTCCAAGCTTCTGGA 1727

Qy 472 rValGlnValValAspProGlyProValLeuArgAspMetAlaPheSerLysAspHisGl 492

Db 1728 TATCTTCGAGTGGTCTCAACAGGAGGATCCAAAGTATGGAATCAGCCAGCACGATAA 1787

Qy 492 uGlnLeuTrpIleMetSerGluArgGlnLeuThrArgValProValGluSerCysGly-- 511

Db 1788 GAGCCCTACATTCGACCGATCGATCGATCGATCAAGCAATCGACCTGCCATGTGCAATCG 1847

Qy 512 -GlnTrpGlnSerCysGlyGluCysLeuGlySerGlyAspProHisCysGlyTrpCysVa 531

Db 1848 CGGTTAGCAACAAGTCTCCGCTGGTGC-----CGTGATCCCTACTGCGGCTGGGATAA 1901

Qy 531 lLeuHisAsnThrCysThrArgLysGluArg----- 541

Db 1902 GGAGGCCAATAGTCCGACCGTACGAGCTGGATTTACTGCAGGATGTGCCAATGAAC 1961

Qy 542 -----CysGluArgSerLysGluProArgPheAlaSerGluMetLysGlnCy 558

Db 1962 GAGTGACATTTGCGATTCGAGTGTCTGAAAGAAAGATTTGTGTGACCTATGCCAGAG 2021

Qy 558 sValArgLeuThrValHisProAsnAsnIleSerValSerGlnTrpAsnValLeuVa 578

Db 2022 TGTATCTCTGGGCTGTTTC----- 2040

Qy 578 lLeuGluThrTrpAsnValProGluLeuSerAlaGlyValAsnCysThrPheGluAspLe 598

Db 2041 -----GTCAAATACCGAAGTGTGGAAGATGAGCAAGTGCACCTGTATCATCA 2090

Qy 598 userGluMetAspGlyLeuValValGlyAsnGlnIleGlnCysTrpSerProAlaAlaLy 618

Db 2091 CTCCAAGGACAGGAGCGGTAC-----GAGATTCGT---TACTCGCGCCACCAA-- 2136

Qy 618 sGluValProArgIleIleThrGluAsnGlyAspHisHisValValGlnLeuGlnLeuLy 638

Db 2137 -----TACATTGAGACCCAGCGGCTGGCTGGTGTGTTCTTCGCGTAA 2180

Qy 638 sSerLysGluThrGlyMetThrPheAlaSerThrSerPheValPheTrpAsnCysSerVa 658

Db 2181 CGAAGCGGATGGTGGTTCG-----TACGATTCGCAATTT 2213

Qy 658 lHisAsnSerCysLeuSerCysValGluSerProTrpArgCysHisTrpCysLysTrpAr 678

Db 2214 GGGCGGCTCGCTTTG-----TGCACGTACAA 2240

Qy 678 g-----HisValCysThrHisAspProLysThrCysSerPheGlnGlu 692

Db 2241 CATTCAGTGGATGCCACAGATGCACCTCCGCCCAACAGAGTAATGACTATCATCAAAA 2298

RESULT 11

PCT-US94-10151A-61

Sequence 61, Application PC/TUS9410151A

GENERAL INFORMATION:

APPLICANT: The Regents of the University of California

TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 66

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR HOBBACH TEST ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/10151A

FILING DATE: 13-SEP-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 27299 FHT UR

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:

LENGTH: 2670 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 268..2439

PCT-US94-10151A-61

Alignment Scores:

Pred. No.: 1.47e-13 Length: 2670

Score: 227.00 Matches: 147

Percent Similarity: 37.22% Conservative: 121

Best Local Similarity: 20.42% Mismatches: 259

Query Match: 2.27% Indels: 194

DB: 5 Gaps: 34

US-09-964-956-13 (1-1896) x PCT-US94-10151A-61 (1-2670)

Qy 53 HisLeuValValAspGluArgThrGlyHisLeuGlyAlaValAsnArgIleTrp 72

Db 481 CATATGAACGAAGATCGAGATACG-----CTCTATGTGGAGCCATGGATCGGTATTC 534

Qy 73 LysLeuSerSerAspLeuLysValLeuValThrHisGluThrGlyProAspGluAspAsn 92

Db 535 CGTGTG-----AACCTCGAGAATATCTCTCATCATCCATTTGTAATCGGGATCGGATCAAC 588

Qy 93 ProLysCysTrpProArgIleValGlnThrCysAsnGluProLeuThrThrAsn 112

Db 589 -----TTGGAGCAACACAGGAGATGATGTGTTAGTCCGTC 624

Qy 113 AsnValAsnLysMetLeuLeuIleAspTrpLysGluAsnArgLeuIleAlaCysGlySer 132

Db 625 TCCAAAGGCAAGATCAGATCTTCGACTGCAAG---AACCATGTGCGGTGTCTATCCACTCA 681

Qy 133 LeuTrpGlnGlyIleCysLysLysLeuLeuArgLeuAspLeuPheLysLeuGlyGluPro 152

Db 682 ATGGACCAAGGG-----GATAGGCTCTATGTATGTCGCGCACCAAC 720



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713D  
FILING DATE: 13-SEP-1993  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4341  
TELEFAX: (415)343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2854 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 451..2640  
S-08-121-713D-57

Alignment Scores:  
Query No.: 2,44e-11 Length: 2854  
Score: 207.00 Matches: 121  
Percent Similarity: 36.96% Conservative: 110  
Best Local Similarity: 19.36% Mismatches: 234  
Query Match: 161 Indels: 161  
Gaps: 30

US-09-964-956-13 (1-1896) x US-08-121-713D-57 (1-2854)

QY	6	TyrSerThrThrCysLeuSerHisLeuMetValGlyMetGlySerSerThrLeu	25
Db	485	TTTGGTGGCGTGCACGGCGCGCAT-----GGGTCAACG-----	520
QY	26	LeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheVal---ThrPheArg	44
Db	521	---ACGTGACGCCCAAGATGTACGTCCAGTTCGGTGAGGAACGG-GTGCACACGCTTCCTG	576
QY	45	GlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIleTyr	64
Db	577	GGCAATGATCGGCACAAAGACCACTTCAACCTGCTGGAGAGGACCACTCGCTCCTC	636
QY	65	LeuGlyAlaValAsnArgIleTyrLysLeuSer---SerAspLeu-----	78
Db	637	GTAGGAGCTAGGAACATCGCTACAATATCAGCCTTCGAGACCTCACAGAATTCACCGAG	696
QY	79	LysValLeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyr---Pro	97
Db	697	CAGAGGATCAGTGGCAGCTCGTCAGGTGCCATCGCGAG-----CTCTGCTACTCTCAAG	750
QY	98	ProArgIleValGlnThrCysAsnGluProLeuThrThrThrAsnAsnValAsnLysMet	117
Db	751	GGGAAGTCAGAGCAGCTGCCAG-----AACTACATCCGAGTCCGTG	792
QY	118	LeuLeuIleAspTyrLysGluAsnArgLeuIleAlaCysGly---SerLeuTyrGlnGly	136
Db	793	CGCAAAATTGAC-----GATGACCGGTACTCATCTCGCGTACGAACGCCCTATAAGCCA	846
QY	137	IleCysLysLeuLeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLys	156
Db	847	CTATGTCGGCACTACGCCCTCAAGAT-----GGAGATTATTTGTAGAGAAA	894
QY	157	Glu-----	164
Db	895	GAATATGAGGGAAGGATTGTGCCCAATTGACCTGACCCACACAGCAGCTGCAATATAC	954

QY	165	GluSerGlySerValPheGlyValIleValSer---TyrSerAsnLeuAspAspLysLeu	183
Db	955	AGTAGGGCAATTTACTCAGCAACAGTGGCAGACTTCTCTGGAAGTCCCTCTCATA	1014
QY	184	PheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSerArgLys	203
Db	1015	TACCGCGGCG-----CCTCTAAGAACACAGAGAGATCT	1044
QY	204	LeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheVal	223
Db	1045	GACCTCAAAACAATTAATGCTCTCAACACATGGAGTACATGATTTTATA	1104
QY	224	AlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAspIleTyr	243
Db	1105	-----TTCTTCTTCTCCGAGAGACTGCTGTTGAG	1134
QY	244	TyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMet	263
Db	1135	TACATC-----AACTCGGAAGGCTATCTAT-----	1161
QY	264	ValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuValArgLeu	283
Db	1162	-----TCAAGAGTTGCCAGAGTC	1179
QY	284	CysLysGluAspThrAla-----PheAsnSerTyrValGlu	295
Db	1180	TGTAACATCACAAAGGCGCCCTCATCAGGGTGGTGACAGATGGACTTCTTTTGGAA	1239
QY	296	ValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuAlaAlaTyrLeu	315
Db	1240	TCACGCTGAACTGTCCTCCGTCCTGGA---GATTATCCATTTACTTCAATGAAATTCAG	1296
QY	316	SerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspLeuLeu	335
Db	1297	TCAACAAGTGACATCATTAAGGAAATATATGTTGTT-----CAAGTGAGAAATCATC	1350
QY	336	PheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeu	355
Db	1351	TACGGTGTCTTCACG-----ACACAGTGAACCTCTATTGGTGGCTCTGCTGTT	1398
QY	356	CysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyr	375
Db	1399	TGTGCTTCAGTATGAAGTCA-----ATACCTGAGTCATTTGATGGTCCATTT	1446
QY	376	ArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIle-----	391
Db	1447	AAAGAGCAGGAACGATGAACCTCAAACTGGTGGCAGTGCCCAAGCCTTAAAGTGCAGAA	1506
QY	392	-----ProCysSerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeu	407
Db	1507	CCAAGGCTGGACATGTCGATGACACAGTCGTACACTTCCTGAT-----GTG	1554
QY	408	AspMetAsnAlaProLeuGlyValSerAspMetValArgGlyIleProValPheThrGlu	427
Db	1555	TCTGTCAATTTTGTAAAGTCACATCATGTGATGAGCGCGCGCCAGCATTTTCTACT	1614
QY	428	AspArgAspArgMetThrSerValIleAlaTyrValTyr-----	440
Db	1615	CGGCAATTCCTCATTCCGATCATGTTACAGTACAGATTTACAAAATACTGTTGATCAA	1674
QY	441	-----LysAsnHisSerLeuAlaPheValGlyThrLysSerGly	453
Db	1675	CAAGTCCGAACACAGATGGGAAAGCGTATGATGCTCTTTTATAGGAACATGATGTCG	1734
QY	454	LysLeuLysLysIle-----ArgValAspGlyPro	463
Db	1735	AAAGTGATAAAAGCTTTGAACCTCTGCTCTTTGATTTCATCTGATCTAGAT-----	1788
QY	464	ArgGlyAsnAlaLeuGlnTyrGluThrValGlnValValAspProGlyProValLeuArg	483
Db	1789	-----AGTGTGTAATAGAAAGTCAAGTGTTCGCCACCTGGAGTACCTGTTAAG	1839
QY	484	AspMet-----AlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGluArg	500

Db 1840 AACCTGATGTGGTGAATGATGGGATAGCAAGCTGGTGTGTCTGATCAT 1899  
Qy 501 GlnLeuThrArgValProValGluSerCysGly-----GlnTyrGlnSerCysGlyGlu 518  
Db 1900 GAGATTCTGGCAATTAAGCTTCACTGTGTGGCTCAGATAAATAACAATAATGTCGAGAA 1959  
Qy 519 CysLeuGlySerGlyAspProHisCysGlyTyrCysValLeuHisAsnThrCysThrArg 538  
Db 1960 TGTGTGTCCTTGCAGATCCTTACTGTGCTGGGATGAGCAATAGAAATTAATAATGATACAGCT 2019  
Qy 539 LysGluArgCysGluArgSerLysGluProArgPhe-----AlaSer 553  
Db 2020 GTAGGTTCAACAGACTGGAGTGTGGAAAAAGCGCTTTATTCAGAACATTTTCACTCGGT 2079  
Qy 554 GluMetLysGlnCys 558  
Db 2080 GAACATAAGCTTGT 2094

RESULT 13  
US-08-835-268-57  
; Sequence 57, Application US/08835268  
; Patent No. 5807826  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S.  
; APPLICANT: Kolodkin, Alex L.  
; APPLICANT: Matthes, David  
; APPLICANT: Bentley, David R.  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/835,268  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/121,713  
; FILING DATE: 13-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B94-002-1  
; TELEPHONE: (415)343-4341  
; TELEFAX: (415) 343-4342  
; TELEX:  
; INFORMATION FOR SEQ. ID NO: 57:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2854 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 451..2640  
US-08-835-268-57

Alignment Scores: 2.44e-11 Length: 2854  
Pred. No.: 207.00 Matches: 121  
Score:

Percent Similarity: 36.96% Conservative: 110  
Best Local Similarity: 19.36% Mismatches: 234  
Query Match: 2.07% Indels: 161  
DB: 1 Gaps: 30  
US-09-964-956-13 (1-1896) x US-08-835-268-57 (1-2854)  
Qy 6 TrpAsnTrpThrCysLeuSerHisLeuLeuMetValGlyMetGlySerSerThrLeu 25  
Db 485 TTTGGTGGCGCTGCACGCCGCCGAT-----GGGTCAACG----- 520  
Qy 26 LeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheVal---ThrPheArg 44  
Db 521 ---ACGTGAGCCCAAGATGTACGTCCAGTTCGGTGAGGAACGG-GTGCACGCTTCCTG 576  
Qy 45 GlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIleTyr 64  
Db 577 GCGAATGAATCGCACAAAGACCTTCAAGCTGCTGGAGAGGACCAACAACCTCGCTCCTC 636  
Qy 65 LeuGlyAlaValAsnArgIleTyrLysLeuSer---SerAspLeu----- 78  
Db 637 GTAGGAGCTAGGAACATCGTCTACAATATACACCTTCGAGACCTCACAGAATTCACCGAG 696  
Qy 79 LysValLeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyr---Pro 97  
Db 697 CAGAGGATCGAGTGGCACTCGTCAAGTGCCTCGCCATCGCGAG-----CTCTGCTACCTCAAG 750  
Qy 98 ProArgIleValGlnThrCysAsnGluProLeuThrThrAsnAsnValAsnLysMet 117  
Db 751 GCGAATGAATCGCACAAAGACCTTCAAGCTGCTGGAGAGGACCAACAACCTCGCTCCTC 792  
Qy 118 LeuLeuIleAspTyrLysGluAsnArgLeuIleAlaCysGly---SerLeuTyrGlnGly 136  
Db 793 GCGAATAATTGAC-----GATGACCGCGTACTACTCTCGGTACGACACCCCTATAGCCA 846  
Qy 137 IleCysLysLeuLeuArgLeuGluAspLeuPheLysLeuGlyGluProTyrHisLysLys 156  
Db 347 CTATGTCGCCACTACGCCCTCAAGCAT-----GGAGATTATCTGTAGAGAAA 894  
Qy 157 Glu-----HisTyrLeuSerGlyValAsn 164  
Db 895 GAATATGAGGGAAGAGGATTTGCCCCATTGACCTGACCAACACACGACTCCAAATATAC 954  
Qy 165 GluSerGlySerValPheGlyValIleValSer---TyrSerAsnLeuAspLysLeu 183  
Db 955 AGTGAGGACAATCTGACTCACCACAGTGGCAGACTTCTTGGAACTCACCTCTCAT 1014  
Qy 184 PheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSerArgLys 203  
Db 1015 TACCGCGGC-----CCTCTAAGAACACAGAGATCT 1044  
Qy 204 LeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheVal 223  
Db 1045 CACCTCAACAATAATTAATGCTCCTCAACACAACTGGAGTACATGATTTATTA 1104  
Qy 224 AlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAspIleTyr 243  
Db 1105 -----TTCTTCTTTCGAGAGACTGCTGTGAG 1134  
Qy 244 TyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMet 263  
Db 1135 TACATC-----AAGTGGCGAAGGCTATCTAT----- 1161  
Qy 264 ValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuValArgLeu 283  
Db 1162 -----TCAAGAGTTCACAGATC 1179  
Qy 284 CysLysGluAspThrAla-----PheAsnSerTyrValGlu 295  
Db 1180 TGTAACATGACAGAGGGCGGCGCTCATCAGGGTGTGACAGATGACATCTCTTTTGA 1239  
Qy 296 ValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaTyrLeu 315

1240 TCACGCTGAACTGTCGCTCCCTGGA---CATTATCCATTTTACTTCAATGAATTCAG 1296  
Qy SerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspLeuLeu 335  
1297 TCACAAAGTGCATCATTTGAAGAAATATATGCTGT-----CAAGTGGAGAACTCATC 1350  
Qy PheThrValPheSerLysGlyClnLysArgLysMetLysSerLeuAspGluSerAlaLeu 355  
1351 TACGGTGTCTTTCACG-----ACACGAGTGAATCTATTGGTGGCTCTGCTGTT 1398  
Qy CysIlePheLeuLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysThr 375  
1399 TGTGCTTCATGATCAAGTCA-----ATACCTTGAGTCATTTGATGCTCCATTT 1446  
Qy ArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIle----- 391  
1447 AAAGAGCAGGAACGATCAACTCAACTGCTGGCAGTGCCAAAGCTTAAAGTGCAGAA 1506  
Qy 392 -----ProCysSerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeu 407  
1507 CCAAGGCTGGCAATGTGTGAATGACAGCTCGTACACTTCTGAT-----GTG 1554  
Qy 408 AspMetAsnAlaProLeuGlyValSerAspMetValArgGlyIleProValPheThrGlu 427  
1555 TCTGTCAATTTTGAAGTCAACATACACTGATGATGATGCTCTTTATAGAACTGATGCG 1614  
Qy 428 AspArgAspArgMetThrSerValIleAlaTrpValTy-----ArgValAspGlyPro 463  
1615 CGGCCAATTCATTCGCGATCAGCTTACAGTACAGATTTACAAAATACAGTGTGATCAA 1674  
Qy 441 -----LysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
1675 CAAGTCCGACACAGATGGGAAGCGTATGATGCTCTTTATAGAACTGATGCG 1734  
Qy 454 LysLeuLysLysIle-----ArgValAspGlyPro 463  
1735 AAAGTGAATAAAGCTTTGAACCTCGCTCTTGTGATTCATGATGATGATGAT----- 1788  
Qy 464 ArgGlyAsnAlaLeuGlnTyThrValGlnValValAspProGlyProValLeuArg 483  
1789 -----AGTGTGTAATAGAGAACTGCAAGTGTGCCACCTGGAGTACCTGTAAG 1839  
Qy 484 AspMet-----AlaPheSerLysAspHisGluGlnLeuTyrlleMetSerGluArg 500  
1840 AACCTGTATGTGTGCGAATGATGGGATGATGATGATGATGATGATGATGATGAT 1899  
Qy 501 GlnLeuThrArgValProValGluSerCysGly-----GlnTyThrGlnSerCysGlyGlu 518  
1900 GAGATTCTGGCAATTAAGCTTTCATCGTTGTGGCTCAGATAAATAACAAATTTGTCGAGAA 1959  
Qy 519 CysLeuGlySerGlyAspProHisCysGlyTrpCysValLeuHisAsnThrCysThrArg 538  
1960 TGTGTGCTTGTGCAAGATCCTTACTGTGTCATGGGACCAATGTAGAATTAATAATGTACAGCT 2019  
Qy 539 LysGluArgCysGluArgSerLysGluProArgArgPhe-----AlaSer 553  
2020 GTAGGTTTCAACAGCTGAGTGTGGAAAAAGAGCGTTTATTCAGAACATTTTCACTCGGT 2079  
Qy 554 GluMetLysGlnCys 558  
2080 GAACATAAAGCTTGT 2094

RESULT 14  
US-09-060-692-57  
Sequence 57, Application US/09060692  
Patent No. 5935865  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey S.  
APPLICANT: Kolodkin, Alex L.  
APPLICANT: Matthews, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESS:  
ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/060,692  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4341  
TELEFAX: (415) 343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2854 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 451..2640  
US-09-060-692-57

Alignment Scores:  
Pred. No.: 2,44e-11 Length: 2854  
Score: 207.00 Matches: 121  
Percent Similarity: 36.96% Conservative: 110  
Best Local Similarity: 19.36% Mismatches: 234  
Query Match: 2.07% Indels: 161  
DB: 2 Gaps: 30

US-09-964-956-13 (1-1896) x US-09-060-692-57 (1-2854)

Qy 6 TrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMetGlySerSerThrLeu 25  
Db 485 TTTGGTGGCGCTGCACCCGCCGCT-----GGGTCAACG----- 520  
Qy 26 LeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheVal---ThrPheArg 44  
Db 521 ---ACGTGAGCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 576  
Qy 45 GlyGluProAlaGluGlyPheAsnHisLeuValAspGluArgThrGlyHisIleTy 64  
Db 577 GGCAATGAATCGCACAAAGACCACTTCAAGCTGCTGGAGAGGACCACTCGCTCCTC 636  
Qy 65 LeuGlyAlaValAsnArgIleTyThrLysLeuSer---SerAspLeu----- 78  
Db 637 GTAGGAGCTAGGAACATCGCTTACAATATCAGCTTCGAGACCTCACAGAATTCACCGAG 696  
Qy 79 LysValLeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTy---Pro 97  
Db 697 CAGAGGATCGAGTGGCAGCTCGCAGTGGCCATCGCGAG-----CTCTGCTACCTCAAG 750  
Qy 98 ProArgIleValGlnThrCysAsnGluProLeuThrThrAsnAsnValAsnLysMet 117  
Db 751 GGAAGATCAGAGGACGACTGCCAG-----AACTACATCCGAGTCTCTG 792



; LENGTH: 2854 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 451..2640  
 ; US-09-833-391-57

## Alignment Scores:

Pred. No.: 2,44e-11 Length: 2854  
 Score: 207.00 Matches: 121  
 Percent Similarity: 36.96% Conservative: 110  
 Best Local Similarity: 19.36% Mismatches: 234  
 Query Match: 2.07% Indels: 161  
 DB: 3 Gaps: 30

US-09-964-956-13 (1-1896) x US-08-833-391-57 (1-2854)

QY 6 TrpAsnTrpThrCysLeuLeuSerHisLeuLeuMetValGlyMetGlySerSerThrLeu 25  
 DB 485 TTTGGTGGCGCGTGCACGCCGCCAT-----GGGTCAAG----- 520  
 QY 26 LeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheVal---ThrPheArg 44  
 DB 521 ---ACGTACGCCCAAGATGATCTCCAGTTCGGTGAGGAACGG-GTGCACACGCTCTCTG 576  
 QY 45 GlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIleTyr 64  
 DB 577 GGCAATGAATCGCACAAAGACCACTTCAAGCTGCTGGAGAGGACCACAACTCGCTCCTC 636  
 QY 65 LeuGlyAlaValAsnArgIleTyrLysLeuSer---SerAspLeu----- 78  
 DB 637 GTAGGATCGAGAACATCGCTACAATATCAGCTTCGAGACCTCACAGAATTACCCGAG 696  
 QY 79 LysValLeuValThrHisGluThrGlyProAspGluAspAsnProLysCysTyr---Pro 97  
 DB 697 CAGAGGATCGAGTGGCAGCTCGTCAAGTGCCCATCGCGAG-----CTCTGCTACCTCAAG 750  
 QY 98 ProArgIleValGlnThrCysAsnGluProLeuThrThrAsnAsnValAsnLysMet 117  
 DB 751 GGGAAAGTCAGAGGACGACTGCGAG-----AATACATCGAGCTCTG 792  
 QY 118 LeuLeuIleAspTyrLysGluAsnArgLeuIleAlaCysGly---SerLeuTyrGlnGly 136  
 DB 793 CCGAAATTGAC-----GATGACGCGCTACTCATCTCGGTACGAACGCCCTATAAGCCA 846  
 QY 137 IleCysLysLeuLeuArgLeuGluAspLeuPheLeuLysLeuGlyGluProTyrHisLysLys 156  
 DB 847 CTATGTCGGCACTACGCCCTCAAGGAT-----GGAGATTATGTTGTAGAGAAA 894  
 QY 157 Glu-----HisTyrLeuSerGlyValAsn 164  
 DB 895 GAATATGAGGGAAGAGGATTGCGCCATTGACCCCTGACCACCAACAGCACTCAATATAC 954  
 QY 165 GluSerGlySerValPheGlyValIleValSer---TyrSerAsnLeuAspAspLysLeu 183  
 DB 955 AGTGAGGACCAATTTACTTCAGCAACAGTGGCAGACTTCTCTGGAACCTGACCCCTCTCATA 1014  
 QY 184 PheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSerArgLys 203  
 DB 1015 TACCGCGGC-----CCTCTAGACAGAGAGATCT 1044  
 QY 204 LeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGluPheVal 223  
 DB 1045 GACCTCAACAAATTAATCTCTTCTTGTCAACACAAATGGAGTACAATGATTTTATA 1104  
 QY 224 AlaSerMetIleLysIleProSerAspThrPheThrIleIleProAspPheAspIleTyr 243  
 DB 1105 -----TTCTTCTTCTTCGAGAGACTGCTGTGAG 1134  
 QY 244 TyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnProGluMet 263

DB 1135 TACATC-----AACTCGGAAAGCGCTATCAT----- 1161  
 QY 264 ValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuValArgLeu 283  
 DB 1162 -----TCAAGAGTTGCCAGAGTC 1179  
 QY 284 CysLysGluAspThrAla-----PheAsnSerTyrValGlu 295  
 DB 1180 TGTAAACATGACAAGGGCGGCCCTCATCAGGGTGTGTGACAGATGGACTTCTTTTGA 1239  
 QY 296 ValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAlaTyrLeu 315  
 DB 1240 TCACGTCTGAACTGTTCGGTCCCTGGA---GATATCCATTTTACTTCAATCAATTCAG 1296  
 QY 316 SerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspLeu 335  
 DB 1297 TCACAAAGTGACATCATTTGAAGAAATTTATGCTGCT-----CAAGTGGAGAACTCATC 1350  
 QY 336 PheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeu 355  
 DB 1351 TACGTGTCTTTCACG-----ACACCAGTGAACCTTATTTGGTGGCTCTGCTGTT 1398  
 QY 356 CysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSerCysTyr 375  
 DB 1399 TGTGCTTTCAGTATGCAAGTCA-----ATACTTGAGTCAATTTGATGCTCCATTT 1446  
 QY 376 ArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIle----- 391  
 DB 1447 AAAGAGCAGAAACGATGAACCTCAAACCTGTTGGCAGTCGCCAGCCCTTAAAGTGCAGAA 1506  
 QY 392 -----ProCysSerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeu 407  
 DB 1507 CCAAGCCCTGGACAATGTGTGAATGACGTGACGTACACTTCCCTGAT-----GTG 1554  
 QY 408 AspMetAsnAlaProLeuGlyValSerAspMetValArgGlyIleProValPheThrGlu 427  
 DB 1555 TCTGTCAATTTGTAAAGTCACATCACATCATGATGATGCTCTGTTATAGCACTGATGATGC 1614  
 QY 428 AspArgAspArgMetThrSerValIleAlaTyrValTyr-----ArgValAspGlyPro 463  
 DB 1735 AAAGTGATAAAAGCTTTGAACCTCTGCTCCTCTTTGATTTCATCTGATAGTAGAT----- 1788  
 QY 464 ArgGlyAsnAlaLeuGlnTyrGluThrValGlnValValAspProGlyProValLeuArg 483  
 DB 1789 -----AGTGTGTAATAGAAGAACTGCAAGTGTGGCCACTGCGAGTACCTGTTAAG 1839  
 QY 484 AspMet-----AlaPheSerLysAspHisGluGlnLeuTyrIleMetSerGluArg 500  
 DB 1840 AACCTGTATGTGGCGCAATGGATGGGATGATAGCAAGCTGTTGTTGCTGCTGATGAT 1899  
 QY 501 GlnLeuThrArgValProValGluSerCysGly-----GlnTyrGlnSerCysGlyGlu 518  
 DB 1900 GAGATTCTGCAATTAAGCTTCATCGTGTGGCTCAGATAAAATAACAAATTTGTCGACAA 1959  
 QY 519 CysLeuGlySerGlyAspProHisCysGlyTrpCysValLeuHisAsnThrCysThrArg 538  
 DB 1960 TGTGTGCTTTCGCAATCTTACTGTGATGGGACAATGTAGAATTAATAATGTACAGCT 2019  
 QY 539 LysGluArgCysGluArgSerLysGluProArgPhe-----AlaSer 553  
 DB 2020 GTAGGTTCAACAGACTGGAGTGGTGGAAAGACGCTTTATTCAGAACATTTCACTCGGT 2079  
 QY 554 GluMetLysGlnCys 558  
 DB 554 GluMetLysGlnCys 558

Db 2080 GAACATAAAGCTTGT 2094

Search completed: June 28, 2003, 16:04:17  
Job time : 456 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 15:21:50 ; Search time 564 Seconds  
(without alignments)  
4990.273 Million cell updates/sec

Title: US-09-964-956-13

Perfect score: 9990

Sequence: 1 MKAMPWNWTCLLSHLLVMGM.....QKLAYKLEQVITLMSLSNKK 1896

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Delop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09964956/runat\_24062003\_124342\_5354/app\_query.fasta\_1.2055  
-DB=PublishedApplications\_NA -QFWT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09964956.cgn\_1.1.376.@runat\_24062003\_124342\_5354  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications\_NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PTC\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PTC\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	6765	67.7	6730	9 US-10-175-523-95 Sequence 95, Appl
2	2580	25.8	6252	10 US-09-964-824A-313 Sequence 313, Appl
3	1372	13.7	2597	9 US-10-245-103-91 Sequence 91, Appl
4	1372	13.7	2597	9 US-10-245-107-91 Sequence 91, Appl

ALIGNMENTS

RESULT 1

US-10-175-523-95

; Sequence 95, Application US/10175523

; Publication No. US20030096264A1

; GENERAL INFORMATION:

; APPLICANT: Brockman, Jeffrey

; APPLICANT: Evans, David

; APPLICANT: Hook, Derek

; APPLICANT: Klimczak, Leszek

; APPLICANT: Laeng, Pascal

; APPLICANT: Palfreyman, Michael

; APPLICANT: Rajan, Prithi

; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)

; FILE REFERENCE: 3235/1J795-US3

; CURRENT APPLICATION NUMBER: US/10/175,523

; CURRENT FILING DATE: 2002-06-18

; PRIOR APPLICATION NUMBER: US 60/299,151

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: US 60/317,828

; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: US 60/325,150

; PRIOR FILING DATE: 2001-09-25

; PRIOR APPLICATION NUMBER: US 60/333,047

; PRIOR FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: US 60/349,936

; PRIOR FILING DATE: 2002-01-18

; PRIOR APPLICATION NUMBER: US 60/361,834

; PRIOR FILING DATE: 2002-03-04





Db	4624	ATCAACAACAAGGTGTTCTTGCTGACCTTCATCCGTACACTCGAAGTACAGCCAGCTTC	4683
QY	1379	SerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeu	1398
Db	4684	TCCATCGGAGACCGTGGGACGTCCTCTCTCATCATGACAGCCCTTCAGGGTCGCCTA	4743
QY	1399	GluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGlu	1418
Db	4744	GAATATGCCACTGATGCTCTCAAGCAGCTGCTCTCTGACCTCATGTGACAGAACTGGAG	4803
QY	1419	SerLysAsnHisProLysLeuLeuLeuArgThrGluSerValAlaGluLysMetLeu	1438
Db	4804	AAACAAGAACCCCAAGCTGCTTCTCCGAGGACTGAGTCTGTGGCGGAGAACTGCTG	4863
QY	1439	ThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeu	1458
Db	4864	ACTAACTGGTTGCTTCTTCTACACAAGTTCCTGAAGGAGTGTGCTGGGGAACCACTC	4923
QY	1459	PheSerLeuPheCysAlaIleLysGlnLeuMetGluLysGlyProIleAspAlaIleThr	1478
Db	4924	TTCATGTATACGTGTGCAATCAAGCAGCAGATGGAAGGCCCCCATTTGACGCTTACT	4983
QY	1479	GlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLys	1498
Db	4984	GGTGAGCCCGATACCTCCCTGAGTGAAGACAAGCTCATCCGCGCAGCAGATGAGTATAAG	5043
QY	1499	ThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLys	1518
Db	5044	ACTCTGATCTCTGAACTGTGCAACCTGACAAATGAGAACAGCCAGAGATCCAGTGAA	5103
QY	1519	IleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLys	1538
Db	5104	GTACTAACTGTGACACCATCACTCACTCAAGTGAAGAGATCCTCGATGCGGTATATAAG	5163
QY	1539	AsnValProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTrpArgGlnGly	1558
Db	5164	AATGTCCCTACTCTCCAGCGCAAGGCTGTGGACATGGATCTGGAGTGGCGCAAGGC	5223
QY	1559	SerGlyAlaArgMetIleLeuGlnAspGluAspIleThrLysIleGluAsnAspTrp	1578
Db	5224	CGGATTGCCAGTGGTGTGACGAGCAGACATTAACCAAAATAGAGGGTGACTGG	5283
QY	1579	LysArgLeuAsnThrLeuAlaHisTyrGlnValProAspGlySerValValAlaLeuVal	1598
Db	5284	AAGCGGTTAACACACTGATGATACCAAGTGTGACAGATCGCTGGTGTCTGTGTT	5343
QY	1599	SerLysGlnValThrAlaTyrAsnAlaValAlaAsnSerThrValSerArgThrSerAla	1618
Db	5344	CCTAAGCAGACCTCTCTCTACAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5403
QY	1619	SerLysTyrGluAsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgThr	1638
Db	5404	AGCAGATGACTCTCTCTTACAGTACACAGGCGAGCCAGACAGCTCCGCTCCCGGTC	5463
QY	1639	ProMetIleThrProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHis	1658
Db	5464	CCCATGATCACCCCAAGCTTGAGAGCGGTGTCAAGGTTGGCATCTGTGGAAGAATCAT	5523
QY	1659	GluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLeu	1678
Db	5524	GACCATGTGCACAGAGAGGGTGTACCGGGGCGAGCAAAATGGTGTCTGAGATCTACTTG	5583
QY	1679	ThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspAspLeuPheGluThr	1698
Db	5584	ACCGGCTTCTAGCCACCAAGGCGACCTCGCAAAATTTGTGACGACTTGTGTGAGACC	5643
QY	1699	IlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAsp	1718
Db	5644	TGTCTCAGCACTGTGCACCGGGTAGTGTCTCTCCCTAGCCATCAAGTACATGTTGAT	5703
QY	1719	PheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrp	1738
Db	5704	TTCCTGGATGAGCAGCAGACACAGATATCCAGCACAGATGTGCGGCACACCTGG	5763
QY	1739	LysSerAsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheVal	1758
Db	5764	AAAAGCAACTCCCTTCCACTTCGTTCTTGGGTGATGTATCATCAAGAACCTCAATTGTA	5823
QY	1759	PheAspIleHisLysAsnSerIleThrAspAlaCysLeuSerValAlaGlnThrPhe	1778
Db	5824	TTTGACATCCACAAGGCGAGCATCACAGATGCTGCTCTCTGTGTAGCCAGACCTTT	5883
QY	1779	MetAspSerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeu	1798
Db	5884	ATGGACTCCTGTTCACATCAGAGCAGCCAGCTAGGCAAGAGCTACCTTCCACACAGCTG	5943
QY	1799	LeuTyrAlaLysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSerAspIle	1818
Db	5944	CTCTATGCCAAGGATATCCCAAGTATTAAGAACTGGTAGAAGATACTATGCAGATATT	6003
QY	1819	GlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluGlnSerArg	1838
Db	6004	GCCAAGCTCCCAAGCATTAGTGACCAAGATATGAATGCCCTACCTCGCGGAGCAGTCCCGC	6063
QY	1839	MethHisMetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGly	1858
Db	6064	CTGCATGCTACAGATTCATATGCTGAGCGCCTCAACGAGATCTACTCATATGTGAGC	6123
QY	1859	LysTyrSerGluGluLeuGlyProLeuAspHisAspGlnCysGlyLysGlnLys	1878
Db	6124	AAAGTACAGTGAGGAGCTATCGGGGCACTAGACAGGATGAACAGGCCCGACGCAACGA	6183
QY	1879	LeuAlaTyrLysLeuGluGlnValIleThrLeuMetSerLeuAspSerAsnLys	1896
Db	6184	CTGGCCTACAGGTGGAGCATCTCATCAAGCCATGTCCATAGAGAGCTGAAAG	6237
RESULT 2			
US-09-964-824A-313			
Sequence 313, Application US/09964824A			
Patent No. US20020102531A1			
GENERAL INFORMATION:			
APPLICANT: Horrigan, Stephen			
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign			
TITLE OF INVENTION: Sets			
FILE REFERENCE: 689290-73			
CURRENT APPLICATION NUMBER: US/09/964, 824A			
CURRENT FILING DATE: 2001-09-27			
PRIOR FILING DATE: 2000-09-28			
PRIOR FILING DATE: 2000-09-28			
PRIOR FILING DATE: 2000-09-28			
PRIOR FILING DATE: 2000-09-28			
NUMBER OF SEQ ID NOS: 583			
SOFTWARE: PatentIn version 3.0			
SEQ ID NO 313			
LENGTH: 6252			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-964-824A-313			
Alignment Scores:			
Pred. No.: 7,21e-270 Length: 6252			
Score: 2580.00 Matches: 654			
Percent Similarity: 50.63% Conservative: 344			
Best Local Similarity: 33.18% Mismatches: 713			
Query Match: 25.83% Indels: 260			
DB: 10 Gaps: 61			
US-09-964-956-13 (1-1896) x US-09-964-824A-313 (1-6252)			
QY	24	ThrLeuLeuThrArgGlnProAlaProLeuSerGlnLysGlnArgSerPheValThrPhe	43
Db	35	ACCCTGCTGGCCCTGCTGGCGCAGGTGCCAGCTTGGAGCCCGCCGCAAGCTGACCTTC	94
QY	44	ArgGlyGluProAlaGluGlyPheAsnHisLeuValValAspGluArgThrGlyHisIle	63

```
Db 95 CCGAGCGAG-AAAAGAGCTGAACACCTGGCTGTGGATGAGCGCTCAGCGGTGGTG 148
Qy 64 TyrLeuGlyAlaValAsnArgIleTyrLysLeuSerSerAspLeuLysValLeuValThr 83
Db 149 TACCTGGGGGGCGGTGATCCCTCTACCAAGCTGGATGGGAAGCTGCAGCTGGAGCAGCAG 208
Qy 84 HisGluThrGlyProAspGluAspAsnProLysCysTyrProProArgIleValGlnThr 103
Db 209 GTGGCCACGGCGCCCGCTGGACACAAGAAGTGCACGCCGCCATCAGAGCCAGCCAG 268
Qy 104 CysAsnGluProLeuThrThrAsnAsnValAsnLysMetLeuLeuLeuLeuAspTyrLys 123
Db 269 TCCCATGAG---GCTGAGATGACTGCAATGTCACACAGCTGCTGCTCGACCTCCC 325
Qy 124 GluAsnArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 143
Db 326 AGGAAGCGCTGGTGGAGTGGCGGACCTCTTCAAGGGCATCTGCGCTCTGCGCGCCCTG 385
Qy 144 GluAsp-----LeuPheLysLeuGlyGluProTyrHisLysLysGluHisTyrLeuSer 161
Db 386 AGCAACATCTCCCTCGCGCTGTCTACGAGGACGGCGGAGGAAGTCTTTCGTGGCC 445
Qy 162 GlyValAsnGluSerClySerValPheGlyValIleValSerTyrSerAsnLeuAspAsp 181
Db 446 AGCAATGATGAGGGCGGTGGCCACAGTGGGGCTGGTGAGCTCCACGGGTCTCTGTGTGAC 505
Qy 182 LysLeuPheIleAlaThrAlaValAspGlyLysProGluTyrPheProThrIleSerSer 201
Db 506 CCGGTGCTTTGTGGGCAAGCAATGGGCGCCACAGCAACGGCATCATCTGTGAGCACT 565
Qy 202 ArgLysLeuThrLysAsnSerGluAlaAspGlyMetPheAlaTyrValPheHisAspGlu 221
Db 566 CGGTGTTGGACCGGACTCACAGCAGGAGCGCTTTGAAGCTTACAGGACCAACGCCACC 625
Qy 222 PheValAlaSerMetIleLysLysIleProSerAspThrPheThrIleLeuProAspPheAsp 241
Db 626 TACAAGCGCGGTACCTGTCCACCAACACACAGCAGTTC-----664
Qy 242 IleTyrTyrValTyrGlyPheSerSerGlyAsnPheValTyrPheLeuThrLeuGlnPro 261
Db 665 -----GTGGCGCCTTCGAGGACGGCCCTACGTCTCTTGTCTTCAACACAGCAG 715
Qy 262 GluMetValSerProGlySerThrThrLysGluGlnValTyrThrSerLysLeuVal 281
Db 716 GACAAG---CACCGCGCCCGGAACCCGACG-----CTGCTGGCA 751
Qy 282 ArgLeuLysLysGluAspThrAlaPheAsnSerTyrTyrValGluValProIleGlyCysGlu 301
Db 752 CGCATGTGCAGAGAGACCCCACTACTCTCTACCTGGAGATGGACCTGCAGTGCCCG 811
Qy 302 ArgSerGlyValIleTyrArgLeuLeuGlnAlaAlaTyrLeuSerLysAlaGlyAlaVal 321
Db 812 GACCCCGACATCCAC-----GCCGTGCGCTTT-----GCCACCTGC 847
Qy 322 LeuGlyArgThrLeuGlyValHisProAspAspLeuLeuPheThrValPheSerLys 341
Db 848 CTGGCGCGCTCGCTGGCTGCGCTGCTGGCAGGGTGCTATATGCTGCTTTCAGCAGA 907
Qy 342 GlyGlnLysArgLysMetLysSerLeuAspGluSerAlaLeuCysIlePheIleLeuLys 361
Db 908 GACAGC-----CGGAGCAGTGGGGGCGCGGTGGCGGCGCTCTGCGTGTTCGCGCTGGAC 961
Qy 362 GlnIleAsnAspArgIleLysGluArgLeuGlnSerCystyrArgGlyGluGlyThrLeu 381
Db 962 AGGTGCAGCCCAAGATGAGGAGCCACCGACCGACCTGTTAC-----ACAGACCCCGG 1015
Qy 382 AspLeuAlaTrpLeuLysValLys-----AspIleProCysSerSerAlaLeu 397
Db 1016 GAGGCCGTGACATCTTCTACAAGCCCTTCCACGGCGATATCCAGTGGCGGCCCGCGG 1075
Qy 398 LeuThrIleAspAsnPhe---CysGlyLeuAsp---MetAsnAlaProLeuGlyVal 415
Db 2090 ---ACCCTGAAGGGTTCTCCTGCACGTGGGCGAGTGAGCTTCAAGTTTCAAGTGGAGCGG 2146
```

QY 759 SerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluIleAsnAsn----- 776  
Db 2147 GTGACCATGACGAATCTGGACCTTGGCCTTCGGACCCCAAGCTGTCCCGCATGGCC 2206  
QY 777 -----LeuProValGluLeuThrValValTirPAsnGlyHisPheAsnIleAspAsn 793  
Db 2207 AACGACAGCTGCCCTGACCTCTAGCTCAAGTCTTACGGCAAG---AATATCGAC--- 2260  
QY 794 ProAlaGlnAsnLysValHis-----LeuTyrLysCysGlyAlaMetArgLysSerCys 811  
Db 2261 -----AGCAAGCTCCATGTGACCTCTACAACTGCTCTTGGCCGCGACGAGTGC 2311  
QY 812 GlyLeuCysLeuLysAlaAspProAspPheAlaCysGlyTirPAsnGlnGlyProGlyGln 831  
Db 2312 AGCTGTGCGCGCGCTACCCGACTACCGTGTGCGTGTGCGGGGCCAGCAGG 2371  
QY 832 CysThrLeuArgGlnHisCysProAlaGlnGluSerGlnTirPLeuGluLeuSerGlyAla 851  
Db 2372 TGCCTGTATGAGGCCCTGTGC-----AACACC 2398  
QY 852 LysSerLysCysThrAsnProArgIleThrGluIleIleProValThrGlyProArgGlu 871  
Db 2399 ACCTCCAGTGCCTGCGCGCGCTCATCACCAGGATCAGCGCTGAGACGGCGCCCTGGGT 2458  
QY 872 GlyGlyThrLysValThrIleArgGlyGluAsnLeuGlyLeuGluPheArgAspIleAla 891  
Db 2459 GGGGGCATCCGCATCACCCTCTGGGTTCCTCAATTTGGCGTCCAAAGCAGGGGACATC--- 2515  
QY 892 SerHisValLysValAlaGlyValGluCysSerProLeuValAspGlyTyrIleProAla 911  
Db 2516 CAGAGGATCTCTGTGCGCGCGGAACTGCTCTCTTTCAGCCGGAACGTTACTCGTGTCC 2575  
QY 912 GluGlnIleValCysGluMetGlyGluAlaLysProSerGlnHisAlaGlyPheValGlu 931  
Db 2576 ACCCGGATCGGTGTGTGATC---GAGGCTCGGAGACGCCCTTTCACCGGGGGTGTGAG 2632  
QY 932 IleCysValAlaValCysArgProGluPhe-----MetAlaArgSerSer----- 946  
Db 2633 GTGGACGTC-----TTGCGGAACCTGGCGCGTTCCTCCCAATGTC 2674  
QY 947 GlnLeuTyrTyrPheMetThrLeuThrLeuSerAspLeuLysProSerArgGlyProMet 966  
Db 2675 CAGTTCACTCCCAACAGCCCAAGCCTCTCAGT---GTGGACCGCAGCAGGAGCGGAG 2731  
QY 967 SerGlyGlyThrGlnValThrIleThrAsnLeuAsnAlaGlyLysSerAsnValVal 986  
Db 2732 CGGGCGGCACACACATCCAGCGCACCCACCTGGACACGGGCTCCAGGAGGAC 2791  
QY 987 ValMetPheGlyLysGlnProCysLeuPheHisArgSerProSerTyrIleValCys 1006  
Db 2792 GTG-----CGGGTGACCTCAACGGCGTCCCGTGT 2821  
QY 1007 AsnThrThrSerSerAspGluValLeuGluMetLysValSerValGlnValAspArgAla 1026  
Db 2822 AAAGTGACGAAGTTGGGGCGCAGCTCCAGTGTGTACTGGCCCCCAGCGACACGGGC 2881  
QY 1027 LysIle-----HisGlnAspLeuValPhe 1034  
Db 2882 CAGATGCTTCTGGAGTCTCTACGGGGGTCCCGCGTCCCAACCCCGGCATCTCTTC 2941  
QY 1035 GlnTyrValGluAspProThrIleValArgIleGluProGluTirPserIleValSerGly 1054  
Db 2942 ACCTACCGCGAAACCCGCTACTGCGAGCCTTTCGAGCGGCTACGAAAGCTTTCGCGAGTGT 3001  
QY 1055 AsnThrProIleAlaValTirPglyThrHisLeuAspLeuIleGlnAsnProGlnIleArg 1074  
Db 3002 GCCGCCACATCAACGTCACGGGTTCAGGCTTCAGCTGATCCAGAGGTTTCCCATG--- 3058  
QY 1075 AlaLysHisGlyGlyLysGluHisIleAsnIleCysGluValLeuAsnAlaThrGluMet 1094  
Db 3059 -----GTGGTTCATCGGGAGGCCCTGTCAGTCTCGTGGCAGCG 3094

QY 1095 ThrCysGlnAla-----ProAlaLeuAlaLeuGlyProAspHis----- 1107  
Db 3095 CCGCGGAGGCTGAATCCTCGACGCCCATGAGGTGGGTACAGACTACGAGTGTTCAC 3154  
QY 1108 -----GlnSerAspLeuThrGluArgProGluGluPheGly 1119  
Db 3155 AATGACACCAAGGTCTCTCTGCTCCCGCGTGTCCCTGAGGAGCAGCAGGCTACAAC 3214  
QY 1120 PheIle-----LeuAspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsn 1135  
Db 3215 CTCAGGTGCTGATGATGACGGGACCGTGCCTCTCAGACAGAGCGCGGCC 3274  
QY 1136 PheThrTyrTyrProAsnProValPheGluAlaPheGlyProSerGlyIleLeuGluLeu 1155  
Db 3275 TTCGAGTACGTGCTGACCCACCTTTGAGAACTTC----- 3310  
QY 1156 LysProGlyThrProIleIleLeuLysGlyLysAsnLeuIleProValAlaGlyGly 1175  
Db 3311 -----ACAGGTGGCGTCAAGAGCAGGTCAACAAGCTATCCACGCCCGGGGCC 3361  
QY 1176 AsnValLysLeuAsnTyrThrVal-----LeuValGlyGluLysProCys 1190  
Db 3362 ATCTGAACAAGCGATGACGTGCGAGGCGCGAGCTTCGTGGTGCAGCGCTGC 3421  
QY 1191 ThrVal-----ThrValSerAspValGlnLeuLeuCysGluSerProAsnLeuIleGlyArg 1209  
Db 3422 ACCATGAAGACGCTGACGGAGACCGACCTGTACTGTGAGCCCCCGGAGGTGACGCCCG 3481  
QY 1210 HisLys-----ValMetAlaArgVal 1216  
Db 3482 CCCAAGCGCGCAGAAACGACGACACACACACTGCCCGAGTTCATTGTCAAGTTC 3541  
QY 1217 GlyGlyMetGluTyrSerProGlyMetValTyrIle-----AlaProAspSerPro 1233  
Db 3542 GCTCTCGCGAGTGTGGCGCGTGGAGTACGACACAGCGGTGTGACGCGCTCGC 3601  
QY 1234 LeuSerLeu-----ProAlaIleValSerIleAlaValAlaGly 1246  
Db 3602 CTCAGCTCATCTGCGCGTGTGTCGTCCTGTCGTCATCGCGGTGTCT--- 3658  
QY 1247 GlyLeuLeuIlePheIleValAlaValLeuIleAlaTyrLysArgLysSerArgGlu 1266  
Db 3659 -----GTCTACTGCTACTTGGAGGAAGACGACGAC 3688  
QY 1267 SerAspLeuThrLeuLysArgLeuGlnMetAsnAsnLeuGluSerArgValAla 1286  
Db 3689 GCCGAACGAGATGATGAGAGATCAAGTCCCGCTGGAGGCGCTGGAGGAGCGTGGC 3748  
QY 1287 LeuGluCysLysGluAlaPheAlaGluLeuThrAspIleHisGlnLeuThrSerAsp 1306  
Db 3749 GACCGTGCAGGAAGAAATTCACAGCTGATCGATGAGGAGGACGACACGAC 3808  
QY 1307 LeuAspGlyAlaGlyIleProPheLeuAspTyrArgThrTyrThrMetArgValLeuPhe 1326  
Db 3809 GTGCACGAGCGCGCATCCCGTGTGACTACAAGACCTACACCGACCGCTCTTCTTC 3868  
QY 1327 -----ProGlyIleGluAspHisProValLeuArgAspLeuGluValProGly 1342  
Db 3869 CTGCCCTCCAAGACGCGACAGGAGCGTATCATCCGCAAGCTGGACATCCCTGAG 3928  
QY 1343 TyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIleAsnAsnLys 1362  
Db 3929 CCGCGCGCGCGTGGTGGAGCGGCCCTCTACAGTCTCTCAACCTCTCAACAGCAAG 3988  
QY 1363 ValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPheSerMetArgAsp 1382  
Db 3989 TCTTCTCATCAATTTTCATCCACCCCTGGAGAACCGGGAGTCTCGCGCGCGGCC 4048  
QY 1383 ArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeuGluTyrAlaThr 1402  
Db 4049 AAGTCTACTTCCGCTCCCTGCTGACGCGTGTGACGGAACTGAGTACTACAG 4108  
QY 1403 AspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGluSerLysAsnHis 1422



; ORGANISM: Homo Sapien	
US-10-245-103-91	
Alignment Scores:	
Pred. No.:	1.1e-138
Score:	1372.00
Percent Similarity:	100.00%
Best Local Similarity:	99.62%
Query Match:	13.73%
DB:	9
US-09-964-956-13 (1-1896) x US-10-245-103-91 (1-2597)	
Qy	194 GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
Db	2 GAGTATTTTCCACCATCTCCAGCGGAACTGACCAAGAACTCTGAGCGGGATGCGATG 61
Qy	214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233
Db	62 TTCGGTACGCTTCCATGATGAGTTCGTGGCTCGATGATTAGATCCCTTCGGACACC 121
Qy	234 PheThrIleLeuProAspPheAspIleTyrValTyrGlyPheSerSerGlyAsnPhe 253
Db	122 TTCACCATCATCCCTGACTTTGATATCTACTATGTTATGTTTTCAGCATGCGCAACTTT 181
Qy	254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
Db	182 GTCTACTTTTGGACCTCCAACTCGAGATGGTGTCTCCACGAGGCTCCACCAAGGAG 241
Qy	274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293
GENERAL INFORMATION:	
; APPLICANT: Baker, Kevin	
; APPLICANT: Eaton, Dan	
; APPLICANT: Filvaroff, Ellen	
; APPLICANT: Goddard, Audrey	
; APPLICANT: Grimaldi, J. Christopher	
; APPLICANT: Gurney, Austin	
; APPLICANT: Smith, Victoria	
; APPLICANT: Stephan, Jean-Phillippe	
; APPLICANT: Watanabe, Colin	
; APPLICANT: Wood, William	
; APPLICANT: Zhang, Zemin	
; APPLICANT: Fong, Sherman	
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC	
; FILE REFERENCE: P3630R1C71	
; CURRENT APPLICATION NUMBER: US/10/245,107	
; PRIORITY FILING DATE: 2002-09-16	
; PRIORITY FILING DATE: 2002-07-18	
; PRIORITY FILING DATE: 2002-07-18	
; PRIORITY FILING DATE: 1997-09-17	
; PRIORITY FILING DATE: 1997-10-24	
; PRIORITY FILING DATE: 1997-11-10	
; PRIORITY FILING DATE: 1998-03-27	
; PRIORITY FILING DATE: 1998-03-27	
; PRIORITY FILING DATE: 1998-05-22	
; PRIORITY FILING DATE: 1998-06-02	
; PRIORITY FILING DATE: 1998-06-18	
; PRIORITY FILING DATE: 1998-06-18	
; PRIORITY FILING DATE: 1998-06-24	
; PRIORITY FILING DATE: 1998-06-25	
; Remaining Prior Application data removed - See File Wrapper or PALM.	
; NUMBER OF SEQ ID NOS: 116	
; SEQ ID NO 91	
; LENGTH: 2597	
; TYPE: DNA	
; ORGANISM: Homo Sapien	
US-10-245-107-91	
Alignment Scores:	
Pred. No.:	1.1e-138
Score:	1372.00
Percent Similarity:	100.00%
Best Local Similarity:	99.62%
Query Match:	13.73%
DB:	9
US-09-964-956-13 (1-1896) x US-10-245-107-91 (1-2597)	
Qy	194 GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
Db	2 GAGTATTTTCCACCATCTCCAGCGGAACTGACCAAGAACTCTGAGCGGGATGCGATG 61
Qy	214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233
Db	62 TTCGGTACGCTTCCATGATGAGTTCGTGGCTCGATGATTAGATCCCTTCGGACACC 121
Qy	234 PheThrIleLeuProAspPheAspIleTyrValTyrGlyPheSerSerGlyAsnPhe 253
Db	122 TTCACCATCATCCCTGACTTTGATATCTACTATGTTATGTTTTCAGCATGCGCAACTTT 181
Qy	254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
Db	182 GTCTACTTTTGGACCTCCAACTCGAGATGGTGTCTCCACGAGGCTCCACCAAGGAG 241
Qy	274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293
Db	242 CAGGTGTATACATCCAACTCGTGGAGCTTTGCAAGAGGAGACACAGCCTTCAACTCCTAT 301
Qy	294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuGlnAlaAla 313
Db	302 GTAGAGGTGCCATTTGGTGTGAGCGCAGTGGGGTGGAGTACCGCTGTCTGACGGCTGCC 361
Qy	314 TyrLeuSerLysAlaGlyValLeuGluGlyArgThrLeuGlyValHisProAspAsp 333
Db	362 TACCTGTCCAAAGCGGGCGCTGTGTCAGGACCTTTGGAGTCCATCCAGATGATGAC 421
Qy	334 LeuLeuPheThrValPheSerLysGlnLysArgLysMetLysSerLeuAspGluSer 353
Db	422 CTGCTCTTTCCAGCTCTTCCAAAGCGGCGAGAGCGGAAATGAAATCCCTGGATGAGTCG 481
Qy	354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
Db	482 GCCCTGTGCATCTTCTATCTTGAAGCAGATAATAGCCGCATTAAAGGAGCGGTGCACTCT 541
Qy	374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
Db	542 TGTACCGGGGGAGGGCAGCTGGACCTGGCTGGCTCAAGGTGAAGGACATCCCTCTGC 601
Qy	394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413
Db	602 AGCAGTGGCTCTTAACCATTCACGATAACTTCTGTGGCTTGGACATGATGCTCCCTCG 661
Qy	414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433
Db	662 GGAGTGTCCGACATGGTGGTGGAAATCCCGCTTTCACGGAGGACAGGACCCATGACG 721
Qy	434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453
Db	722 TCTGTATCGCATGATGCTTACAAGAACCACTCTCTGGCCCTTTGTGGGCGACCAAAAGTGGC 781
Qy	454 LysLeuLysLysIle 458
Db	782 AAGCTAAGAAGGTG 796
RESULT 4	
US-10-245-107-91	
; Sequence 91, Application US/10245107	
; Publication No. US20030068779A1	

```
Db 242 CAGGTGTATACATCAAGCTCTGTAGGCTTTCAGAGGAGGACACACCTTCAACTCCTAT 301
Qy 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313
Db 302 GTAGAGTGCCCATTTGGCTGTGAGGCAGTGGGTGGAGTACCGCTGTGTGAGGCTGCC 361
Qy 314 TyrLeuSerLysAlaGlyAlaValLeuGluArgThrLeuGlyValHisProAspAsp 333
Db 362 TACCTGTCTCAAGAGCGGGCCGTGCTTGGCAGGACCCCTTGGAGTCCATCCAGATGATGAC 421
Qy 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353
Db 422 CTGCTGTCTTACCGCTTCTTCAAGGCGCCAGAGCGGAAATGAAATCCCTGGATGAGTCG 481
Qy 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
Db 482 GCCCTGTGCATCTTTCATCTTGAAGCAGATAATGACCGCATTAAGAGCGGTGCGAGTCT 541
Qy 374 CystyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
Db 542 TGTTCAGCGGGCGAGGCGACGCTGGACCTGGCTGGCTCAAGGTGAAGGACATCCCTGTC 601
Qy 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413
Db 602 AGCATGTGGCTCTTAACCATTCAGCATAACTTCTGTGGCTGGACATGAATGCTCCCTG 661
Qy 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433
Db 662 GGAGTGTCCGACATGGTGGCTGGATTCCTCCGTCCTCCAGGAGGACAGGACCGCATGACG 721
Qy 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453
Db 722 TCTGTATCGCATATGTCTACAAGAACCACTCTCTGGCCCTTTGTGGCCACCAAAAGTGC 781
Qy 454 LysLeuLysLysIle 458
Db 782 AAGCTGAAGAGGTG 796

RESULT 5
US-10-245-143-91
; Sequence 91, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanbe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P36301C90
; CURRENT APPLICATION NUMBER: US/10/245,143
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
```

```
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 91
; LENGTH: 2597
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-143-91

Alignment Scores:
Pred. No.: 1,1e-138 Length: 2597
Score: 1372.00 Matches: 264
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 0
Query Match: 13.73% Indels: 0
DB: Gaps: 0

US-09-964-956-13 (1-1896) x US-10-245-143-91 (1-2597)
Qy 194 GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
Db 2 GAGTATTTTCCACCATCTCCAGCCGGGAAACTGACCAAGAACTCTGAGCGGGATGGCATG 61
Qy 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233
Db 62 TTCGCGTACGCTCTCCATGATGAGTTCGTGCGCTCGATGATTAAAGTATCCCTTCGCACACC 121
Qy 234 PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253
Db 122 TTCACCATCATCCCTGACTTTGATATCTACTATGTCTATGTTTAGCAGTGGCAACTTT 181
Qy 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
Db 182 GTCTACTTTTGGACCTTCCACCTGAGATGGTGTCTCCACCAGGCTCCACCACCAAGGAG 241
Qy 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293
Db 242 CAGGTGTATACATCCAAAGCTCGTGAGGCTTTGCAAGGAGGACACAGCCCTTCAACTCCTAT 301
Qy 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313
Db 302 GTAGAGTGCCCATTTGGCTGTGAGCGGAGTGGGTGGAGTACCGGCTCTCGAGGCTGCC 361
Qy 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333
Db 362 TACCTGTCCAAAGCGGGGGCGCTCTTGGCAGGACCCCTTGGAGTCCATCCAGATGATGAC 421
Qy 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353
Db 422 CTGCTGTCTTACCGCTTCTTCAAGGCGCAGAGCGGAAATGAAATCCCTGGATGAGTCG 481
Qy 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
Db 482 GCCCTGTGCATCTTTCATCTTGAAGCAGATAATGACCGCATTAAGAGGCGGTGCGAGTCT 541
Qy 374 CystyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
Db 542 TGTTCAGCGGGCGAGGCGACGCTGGACCTGGCTGGCTCAAGGTGAAGGACATCCCTGTC 601
Qy 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413
Db 602 AGCATGTGGCTCTTAACCATTCAGCATAACTTCTGTGGCTGGACATGAATGCTCCCTG 661
Qy 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433
Db 662 GGAGTGTCCGACATGGTGGCTGGATTCCTCCGTCCTCCAGGAGGACAGGACCGCATGACG 721
```

QY	434	SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly	453	
Db	722	TCTGTCATCGCATATGCTACAAAGAACCACTCTCTGGCCCTTTGTGGCGACCAAAAGTGGC	781	
QY	454	LysLeuLysLysIle	458	
Db	782	AACTGAAGAAGGTG	796	
RESULT 6				
US-10-245-771-91				
; Sequence 91, Application US/10245771				
; Publication No. US20030068782A1				
; GENERAL INFORMATION:				
; APPLICANT: Baker, Kevin				
; APPLICANT: Eaton, Dan				
; APPLICANT: Filvaroff, Ellen				
; APPLICANT: Goddard, Audrey				
; APPLICANT: Grimaldi, J. Christopher				
; APPLICANT: Gurney, Austin				
; APPLICANT: Smith, Victoria				
; APPLICANT: Stephan, Jean-Phillippe				
; APPLICANT: Watanabe, Colin				
; APPLICANT: Wood, William				
; APPLICANT: Zhang, Zemin				
; APPLICANT: Fong, Sherman				
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC				
; FILE OF INVENTION: ACIDS ENCODING THE SAME				
; FILE REFERENCE: P3630R1C98				
; CURRENT APPLICATION NUMBER: US/10/245,771				
; CURRENT FILING DATE: 2002-09-16				
; PRIORITY APPLICATION NUMBER: 10/197942				
; PRIORITY FILING DATE: 2002-07-18				
; PRIORITY APPLICATION NUMBER: 60/059114				
; PRIORITY FILING DATE: 1997-09-17				
; PRIORITY APPLICATION NUMBER: 60/063046				
; PRIORITY FILING DATE: 1997-10-24				
; PRIORITY APPLICATION NUMBER: 60/065027				
; PRIORITY FILING DATE: 1997-11-10				
; PRIORITY APPLICATION NUMBER: 60/079689				
; PRIORITY FILING DATE: 1998-03-27				
; PRIORITY APPLICATION NUMBER: 60/086478				
; PRIORITY FILING DATE: 1998-05-22				
; PRIORITY APPLICATION NUMBER: 60/087607				
; PRIORITY FILING DATE: 1998-06-02				
; PRIORITY APPLICATION NUMBER: 60/089801				
; PRIORITY FILING DATE: 1998-06-18				
; PRIORITY APPLICATION NUMBER: 60/090557				
; PRIORITY FILING DATE: 1998-06-24				
; PRIORITY APPLICATION NUMBER: 60/090689				
; PRIORITY FILING DATE: 1998-06-25				
; Remaining Prior Application data removed - See File Wrapper or PALM.				
; NUMBER OF SEQ ID NOS: 116				
; SEQ ID NO 91				
; LENGTH: 2597				
; TYPE: DNA				
; ORGANISM: Homo Sapien				
US-10-245-771-91				
Alignment Scores:				
Pred. No.:	1,1e-138	Length:	2597	
Score:	1372.00	Matches:	264	
Percent Similarity:	100.00%	Conservative:	1	
Best Local Similarity:	99.62%	Mismatches:	0	
Query Match:	13.73%	Indels:	0	
DB:	9	Gaps:	0	
US-09-964-956-13 (1-1896) x US-10-245-771-91 (1-2597)				
QY	194	GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet	213	
Db	2	GAGTATTTCCCACTCTCCAGCGGAAACTGACCAAGAAGTCTGTAGGCGGATGGCATG	61	

```
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 91
; LENGTH: 2597
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-851-91

Alignment Scores:
Pred. No.: 11e-138 Length: 2597
Score: 1372.00 Matches: 264
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 0
Query Match: 13.73% Indels: 0
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x US-10-245-851-91 (1-2597)
QY 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
DB 2 GAGTATTTCCACCATCTCCAGCCGGAACACTACCAAGAACTCTGAGCGCGATGGCATG 61
QY 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233
DB 62 TTCGCGTACGCTCTCCATGATGATGTCGCGCTCGATGATTAAGATCCCTTCGGACACC 121
QY 234 PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253
DB 122 TTCACCATCATCCCTGACTTGATATCTACTATGTCTATGTTTACGACGTGGCAACTTT 181
QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
DB 182 GTCTACTTTTGACCTCCACCTGAGATGCTGTCTCCACGAGCTCCACCAAGAGAG 241
QY 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293
DB 242 CAGGTGTATACATCCCAAGCTCGTGAGCGCTTTGCAAGGAGGACACAGCCTTCAACTCCTAT 301
QY 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313
DB 302 GTAGAGGTGCCCATTTGGCTGTGAGCCAGCGAGGTGAGTACCGCTCGCTGCAGGCTGCC 361
QY 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspAsp 333
DB 362 TACCTGTCCAAAGCGGGCGCTGCTTGGCAGAGACCTTGGAGTCCATCCAGATGATGAC 421
QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353
DB 422 CTGCTCTTCAACGCTCTCTCCAAAGGCGCAGAAAGGAAATGAAATCCCTGGATGAGTCG 481
QY 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
DB 116

482 GCCCTGTGCATCTTTCATCTTGAAGCAGATAAATGACCGCATTAAGGAGCGGTGCAGTCT 541
QY 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTyrLeuLysValLysAspIleProCys 393
DB 542 TGTACCAGGCGGAGGAGGACGCTGGACCTGGCTGGCTCAAGTGAAGGACATCCCTCC 601
QY 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413
DB 602 AGCAGTGGCTCTTAAACCATTAACCATTAACCATTAACCATTAACCATTAACCATTA 661
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433
DB 662 GGAGTGTCCGACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 721
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453
DB 722 TCTGTATCGCATATGCTGTACAGAAACCATCTCTCTGGCTTTTGGGCGCCTTTGG 781
QY 454 LysLeuLysLysIle 458
DB 782 AAGCTGAAGAAGGTG 796

RESULT 8
US-10-245-883-91
; Sequence 91, Application US/10245883
; Publication No. US20030068783A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P36303RIC70
; CURRENT APPLICATION NUMBER: US/10/245,883
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 91
; LENGTH: 2597
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-883-91

Alignment Scores:
```

Pred. No.: 11e-138 Length: 2597  
Score: 1372.00 Matches: 264  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.62% Mismatches: 0  
Query Match: 13.73% Indels: 0  
DB: Gaps: 0

US-09-964-956-13 (1-1896) x US-10-245-883-91 (1-2597)

Qy	194	GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet	213
Db	2	GAGTATTTCACCATCTCCAGCGCAAACTGACCAAGAACTCTGAGCGGATGGCATG	61
Qy	214	PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr	233
Db	62	TTCCGCTACGCTCTCCCATGATGAGTTCGTGGCCCTCGATGATTAAGATCCCTTCGGACACC	121
Qy	234	PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe	253
Db	122	TTCCACATCATCCTGACTTTGATATCTACTATGCTATGCTTTAGCAGTGGCAACTT	181
Qy	254	ValTyrPheLeuThrLeuGlnProGluMetValSerProProGlySerThrThrLysGlu	273
Db	182	GTCTACTATTTTGACCTCCACCTGAGATGCTGCTCCACAGGCTCCACCACCAAGGAG	241
Qy	274	GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr	293
Db	242	CAGGTGTATATCATCAAGCTCTGAGGCTTTCAGAGGAGGACACAGCCTTCAACTCCAT	301
Qy	294	ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuGlnAlaAla	313
Db	302	GTAGAGGTGCCATTGGCTGTGAGCGCAGTGGGTGGAGTACCGCTGCTGCAGGCTGCC	361
Qy	314	TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspAsp	333
Db	362	TACCTGTCCAAAGCGGGCGCTGCTGGCAGGACCTTTGGAGTCCATCCAGATGATGAC	421
Qy	334	LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer	353
Db	422	CTGCTCTTCCAGCTCTTCCAAAGCGCCAGAAAGCGGAAATGAAATCCCTGGATGATCG	481
Qy	354	AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer	373
Db	482	GCCTGTGCTATCTTCTTGAAGCAGATAAATGACCGCATTAAGAGCGGCTGCAGTCT	541
Qy	374	CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys	393
Db	542	TGTTACCGGGGAGGCGGACGCTGACCTGGCTGGCTCAAGTGAAGGACATCCCGCTGC	601
Qy	394	SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu	413
Db	602	AGCAGTGGCTCTTAACCATTCAGCATTAATCTGTGGCTGGACATGAATGCTCCCGTG	661
Qy	414	GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr	433
Db	662	GGAGTGTCCGACATGGTGGTGGATTCCTCCGTTCTCCAGGAGGACAGGACCGCATGACG	721
Qy	434	SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly	453
Db	722	TCTGTATCGCATATGCTACAAAGAACCATCTCTGTGGCTTTGTGGGCAACCAAAAGTGGC	781
Qy	454	LysLeuLysLysIle 458	
Db	782	AAAGCTGAAGAGGTG 796	

RESULT 9

US-10-237-535-91  
; Sequence 91, Application US/10237535  
; Publication No. US20030073188A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen

APPLICANT: Goddard, Audrey  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin  
APPLICANT: Smith, Victoria  
APPLICANT: Stephan, Jean-Philippe  
APPLICANT: Watanabe, Colin  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
APPLICANT: Fong, Sherman  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3630RIC3  
CURRENT APPLICATION NUMBER: US/10/237,535  
CURRENT FILING DATE: 2002-09-06  
PRIOR APPLICATION NUMBER: 10/197942  
PRIOR FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/059114  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/063046  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/065027  
PRIOR FILING DATE: 1997-11-10  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/086478  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090689  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/091358  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/099803  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/106932  
PRIOR FILING DATE: 1998-11-03  
PRIOR APPLICATION NUMBER: 60/115554  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119342  
PRIOR FILING DATE: 1999-02-09  
PRIOR APPLICATION NUMBER: 60/123957  
PRIOR FILING DATE: 1999-03-12  
PRIOR APPLICATION NUMBER: 60/123972  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: 60/127372  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/131271  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/133459  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/135725  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/135729  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/135750  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/138385  
PRIOR FILING DATE: 1999-06-09  
PRIOR APPLICATION NUMBER: 60/140653  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: 60/141037  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 60/144732  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/144758  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/144790

```

; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145228
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146843
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/148188
; PRIOR FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/148513
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/149327
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149395
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/150114
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/151700
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/151734
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/162506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171118
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: 60/179851
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 60/180921
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/187202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/198587
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 60/199614
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 60/206330
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206368
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/209832
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: 60/218371
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 60/222695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/229896
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/230621
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/232887
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/235147
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/261878
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/261910
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/261939
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/262150
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/264395
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/266421
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/267623
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/274399
; PRIOR FILING DATE: 2001-03-09

```

```

; PRIOR APPLICATION NUMBER: 60/280982
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/282129
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/282199
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/290589
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: 09/180997
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: 09/267213
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/802706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/924419
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/927796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/929404
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 09/931836
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/941992
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/946374
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 10/001054
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/081056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 10/119480
; PRIOR FILING DATE: 2002-04-09

```

Alignment Scores:

```

Pred. No.: 1,1e-138 Length: 2597
Score: 1372.00 Matches: 264
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 0
Query Match: 13.73% Indels: 0
DB: 9 Gaps: 0

```

US-09-964-956-13 (1-1896) x US-10-237-535-91 (1-2597)

```

Qy 194 GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
Db 2 GAGTATTTTCCACCACATCTCCAGCGGAAACTGACCAAGAACTCTGAGCGGATGCGATG 61
Qy 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233
Db 62 TCCGCGTACGTCTCCATGATGAGTTCGTGGCCTCGATTAAGATCCCTTCGACACACC 121
Qy 234 PheThrIleIleProAspPheAspIleTyrValTyrGlyPheSerSerGlyAsnPhe 253
Db 122 TTCACCATCATCCCTGACTTTTCATATCTATGTCATGCTTTTAGCAGTGGCAACTTT 181
Qy 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273

```

Db 182 GTCTACTTTTGTGACCTTCCAACTGAGATGGTGTCTCCACGAGCTCCACCAAGGAG 241  
Qy 274 GlnValThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293  
Db 242 CAGGTGTATACATCCAGCTCGTGGAGCTTGGCAAGGAGGACACAGCTTCAACTCCTAT 301  
Qy 294 ValGluValProLeuGlyCysGluArgSerGlyValGluTyrArgLeuGluAlaAla 313  
Db 302 GTAGAGGTGCCATTTGGCTGTGAGCGGAGTGGGTGGAGTACCCGCTGCTCGAGGCTGCC 361  
Qy 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333  
Db 362 TACCTGTCCAAAGCGGGCGCTGCTGGCAGGACCTTGGAGTCCATCCAGATGATGAC 421  
Qy 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353  
Db 422 CTGCTCTTCCAGCTTCTTCCAAAGCGGCGAGAAATGAAATCCCTGGATGAGTGG 481  
Qy 354 AlaLeuCysLeuPheLeuLeuLysGlnLeuAspArgLeuLysGluArgLeuGlnSer 373  
Db 482 GCCCTGTGCATCTTCACTTGAAGCAGATAAATGACCGCATTTAAGGAGCGGCTGCAGTCT 541  
Qy 374 CysTyrArgGlyGluThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393  
Db 542 TGTACCGGGGAGGCGACGCTGGACCTGGCTCAAGTCAAGTCAAGGACATCCCTGCG 601  
Qy 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413  
Db 602 AGCAGTGGCTCTTAAACCATTTGACATTAATGACCGCATTTGCGCTGGATGAATCTCCCTCG 661  
Qy 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433  
Db 662 GGAGTGTCCGACATGGTGGTGGATTTCCCGTCTTCACGGAGGACAGGACCGCATGAGC 721  
Qy 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453  
Db 722 TCTGTCTCATCATATGTCTACAAACCACTCTGTGGCTTTGTGGCCACCAAAAGTGGC 781  
Qy 454 LysLeuLysLysIle 458  
Db 782 AAGCTGAAGAGGTG 796  
RESULT 10  
US-10-238-183-91  
Sequence 91, Application US/10238183  
Publication No. US20030073189A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin  
APPLICANT: Eaton, Dan  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Goddard, Audrey  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin  
APPLICANT: Smith, Victoria  
APPLICANT: Stephan, Jean-Phillippe  
APPLICANT: Watanabe, Colin  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
APPLICANT: Fong, Sherman  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3630RIC11  
CURRENT FILING DATE: 2002-09-09  
PRIORITY APPLICATION NUMBER: US/10/238,183  
PRIORITY FILING DATE: 2002-07-18  
PRIORITY APPLICATION NUMBER: 60/059114  
PRIORITY FILING DATE: 1997-09-17  
PRIORITY APPLICATION NUMBER: 60/063046  
PRIORITY FILING DATE: 1997-10-24  
PRIORITY APPLICATION NUMBER: 60/065027  
PRIORITY FILING DATE: 1997-11-10  
PRIORITY APPLICATION NUMBER: 60/079689

PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/086478  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090689  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/091358  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/099803  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/106932  
PRIOR FILING DATE: 1998-11-03  
PRIOR APPLICATION NUMBER: 60/115554  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119342  
PRIOR FILING DATE: 1999-02-09  
PRIOR APPLICATION NUMBER: 60/123957  
PRIOR FILING DATE: 1999-03-12  
PRIOR APPLICATION NUMBER: 60/123972  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: 60/127372  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/131271  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/133459  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/135725  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/135729  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/135750  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/138385  
PRIOR FILING DATE: 1999-06-09  
PRIOR APPLICATION NUMBER: 60/140653  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: 60/141037  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 60/144732  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/144758  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/144790  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/145228  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/145698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: 60/146222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: 60/146843  
PRIOR FILING DATE: 1999-08-03  
PRIOR APPLICATION NUMBER: 60/148188  
PRIOR FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/148513  
PRIOR FILING DATE: 1999-08-12  
PRIOR APPLICATION NUMBER: 60/149327  
PRIOR FILING DATE: 1999-08-17  
PRIOR APPLICATION NUMBER: 60/149395  
PRIOR FILING DATE: 1999-08-17  
PRIOR APPLICATION NUMBER: 60/150114  
PRIOR FILING DATE: 1999-08-20  
PRIOR APPLICATION NUMBER: 60/151700  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: 60/151734  
PRIOR FILING DATE: 1999-08-31

; PRIOR APPLICATION NUMBER: 60/162506  
 ; PRIOR FILING DATE: 1999-10-29  
 ; PRIOR APPLICATION NUMBER: 60/170262  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 60/177118  
 ; PRIOR FILING DATE: 2000-01-20  
 ; PRIOR APPLICATION NUMBER: 60/179851  
 ; PRIOR FILING DATE: 2000-02-02  
 ; PRIOR APPLICATION NUMBER: 60/180921  
 ; PRIOR FILING DATE: 2000-02-08  
 ; PRIOR APPLICATION NUMBER: 60/187202  
 ; PRIOR FILING DATE: 2000-03-03  
 ; PRIOR APPLICATION NUMBER: 60/198587  
 ; PRIOR FILING DATE: 2000-04-18  
 ; PRIOR APPLICATION NUMBER: 60/199614  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 60/206330  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/206368  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/209832  
 ; PRIOR FILING DATE: 2000-06-05  
 ; PRIOR APPLICATION NUMBER: 60/218371  
 ; PRIOR FILING DATE: 2000-07-13  
 ; PRIOR APPLICATION NUMBER: 60/222695  
 ; PRIOR FILING DATE: 2000-08-02  
 ; PRIOR APPLICATION NUMBER: 60/229896  
 ; PRIOR FILING DATE: 2000-09-01  
 ; PRIOR APPLICATION NUMBER: 60/230621  
 ; PRIOR FILING DATE: 2000-09-05  
 ; PRIOR APPLICATION NUMBER: 60/232887  
 ; PRIOR FILING DATE: 2000-09-15  
 ; PRIOR APPLICATION NUMBER: 60/235147  
 ; PRIOR FILING DATE: 2000-09-22  
 ; PRIOR APPLICATION NUMBER: 60/261878  
 ; PRIOR FILING DATE: 2001-01-12  
 ; PRIOR APPLICATION NUMBER: 60/261910  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/261939  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/262150  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/264395  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 60/266421  
 ; PRIOR FILING DATE: 2001-02-02  
 ; PRIOR APPLICATION NUMBER: 60/267623  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/274399  
 ; PRIOR FILING DATE: 2001-03-09  
 ; PRIOR APPLICATION NUMBER: 60/280982  
 ; PRIOR FILING DATE: 2001-04-03  
 ; PRIOR APPLICATION NUMBER: 60/282129  
 ; PRIOR FILING DATE: 2001-04-04  
 ; PRIOR APPLICATION NUMBER: 60/282199  
 ; PRIOR FILING DATE: 2001-04-04  
 ; PRIOR APPLICATION NUMBER: 60/290589  
 ; PRIOR FILING DATE: 2001-05-09  
 ; PRIOR APPLICATION NUMBER: 09/180997  
 ; PRIOR FILING DATE: 1998-11-19  
 ; PRIOR APPLICATION NUMBER: 09/267213  
 ; PRIOR FILING DATE: 1999-03-12  
 ; PRIOR APPLICATION NUMBER: 09/380137  
 ; PRIOR FILING DATE: 1999-08-25  
 ; PRIOR APPLICATION NUMBER: 09/380138  
 ; PRIOR FILING DATE: 1999-08-25  
 ; PRIOR APPLICATION NUMBER: 09/403297  
 ; PRIOR FILING DATE: 1999-10-18  
 ; PRIOR APPLICATION NUMBER: 09/423741  
 ; PRIOR FILING DATE: 1999-11-10  
 ; PRIOR APPLICATION NUMBER: 09/709238  
 ; PRIOR FILING DATE: 2000-11-08  
 ; PRIOR APPLICATION NUMBER: 09/802706

; PRIOR FILING DATE: 2001-03-09  
 ; PRIOR APPLICATION NUMBER: 09/872035  
 ; PRIOR FILING DATE: 2001-06-01  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 09/924419  
 ; PRIOR FILING DATE: 2001-08-06  
 ; PRIOR APPLICATION NUMBER: 09/927796  
 ; PRIOR FILING DATE: 2001-08-09  
 ; PRIOR APPLICATION NUMBER: 09/929404  
 ; PRIOR FILING DATE: 2001-08-13  
 ; PRIOR APPLICATION NUMBER: 09/931836  
 ; PRIOR FILING DATE: 2001-08-16  
 ; PRIOR APPLICATION NUMBER: 09/941992  
 ; PRIOR FILING DATE: 2001-08-28  
 ; PRIOR APPLICATION NUMBER: 09/946374  
 ; PRIOR FILING DATE: 2001-09-04  
 ; PRIOR APPLICATION NUMBER: 10/001054  
 ; PRIOR FILING DATE: 2001-11-30  
 ; PRIOR APPLICATION NUMBER: 10/052586  
 ; PRIOR FILING DATE: 2002-01-15  
 ; PRIOR APPLICATION NUMBER: 10/081056  
 ; PRIOR FILING DATE: 2002-02-20  
 ; PRIOR APPLICATION NUMBER: 10/119480  
 ; PRIOR FILING DATE: 2002-04-09

Alignment Scores:

Pred. No.: 1,1e-138 Length: 2597  
 Score: 1372.00 Matches: 264  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.63% Mismatches: 0  
 Query Match: 13.73% Indels: 0  
 DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x US-10-238-183-91 (1-2597)

QY 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213  
 |||||  
 Db 2 GAGTATTTTCCACCACCTCTCCAGCGGAACTGACCAAGAACTCTCAGCGGGATGCGATG 61  
 QY 214 PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr 233  
 |||||  
 Db 62 TTCGCGTACGCTTCCATGATGAGTTCGTGGCGCTCGATGATTAAAGATCCCTTCGGACACC 121  
 QY 234 PheThrIleProAspPheAspIleTyrValTyrValTyrGlyPheSerSerGlyAsnPhe 253  
 |||||  
 Db 122 TTCACCATCATCCCTGACCTTTGATATCTACTATGCTATGTTTGGTTTACGAGTGGCACTTT 181  
 QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273  
 |||||  
 Db 182 GTCTACTTTTTCACCTCTCAACCTGAGATGGTGTCTCCACAGGCTCCACCAAGGAG 241  
 QY 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293  
 |||||  
 Db 242 CAGGTGTATACATCCAGCTCGTAGGGCTTTCAGAGGAGGACACAGCTTCAACTCCTAT 301  
 QY 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313  
 |||||  
 Db 302 GTAGAGGTGCCCATTTGGCTGTGAGCGCAGTGGGGTGAGTACCGCTGTGCGAGGTGCC 361  
 QY 314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspAsp 333  
 |||||  
 Db 362 TACCTGTCCAAAGCGGGGCCCTGCTTGGCAGGACCTTTGGAGTCCATCCAGATGATGAC 421  
 QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353  
 |||||  
 Db 422 CTGCTCTTTCACCGCTCTTCTCCAAAGGCCAGAGCGGAAATGAAATCCCTGGATGAGTCG 481  
 QY 354 AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373  
 |||||  
 Db 482 GCCCTGTGCATCTTTCATCTTTGAAGCAGATAAATGACCGCATTAAGAGGCGGTGCGACTCT 541  
 QY 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTyrLeuLysValLysAspIleProCys 393

Best Local Similarity:	99.62%	Mismatches:	0
Query Match:	13.73%	Indels:	0
DB:	9	Gaps:	0
US-09-964-956-13 (1-1896) x US-10-238-283-91 (1-2597)			
Qy	194	GluTyrPheProThrIleSerSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet	213
Db	2	GAGTATTTTCCCACCATCTCCAGCGGAAACTGACCAAGAACTCTGAGGCGGATGGCATG	61
Qy	214	PheAlaTyrValPheHisAspGluPheValAlaSerMetIleLysIleProSerAspThr	233
Db	62	TTCCGCTACGCTTCCTCATGATGAGTTCGTGGCTCGATGATTAAAGATCCCTTCGGACACC	121
Qy	234	PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe	253
Db	122	TTACCACTATCATCCTGACTTTGATATCTACTATGCTGATGCTGTTTTAGCAGTGGCACTTT	181
Qy	254	ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu	273
Db	182	GTCTACTTTTGTACCTCCAACTGAGATGGTGCTCCACAGGCTCCACCAAGGAG	241
Qy	274	GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr	293
Db	242	CAGGTGTATACATCCAAGCTCGTGAGGCTTTGCAAGGAGGACACAGGCTTCACTCTCTAT	301
Qy	294	ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla	313
Db	302	GTAGAGGTGCCATTTGGCTGTGAGCGCAGTGGGTGGAGTACCGCCTCTCTCAGGCTGCC	361
Qy	314	TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAspAsp	333
Db	362	TACCTTGTCAAAGCGGGGGCCGTGCTTGGCAGGACCCCTTGGAGTCCATCCAGATGATGAC	421
Qy	334	LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer	353
Db	422	CTGCTCTTCACGCTCTTCTCCAAGGCCAGAGGAGGAAATCAAAATCCCTGGATCAGTCG	481
Qy	354	AlaLeuCysIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer	373
Db	482	GGCCTGTGCATCTTCATCTTTGAAGCAGATAAATGACCGCATTTAAGGAGCGGCTGCAGTCT	541
Qy	374	CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys	393
Db	542	TGTTACCGGGGCGAGGGACGCTGGACCTGGCTGGCTCAAGTCAAGGACATCCCTCCCTGC	601
Qy	394	SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu	413
Db	602	AGCAGTGGCTCTTTAAACCATTTGACGATACTTCTGTGGCCTGGAGATGAATGCTCCCTTG	661
Qy	414	GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr	433
Db	662	GGAGTGTCCGACATGGTGGTGGAAATTCCTGCTTCACGGAGGACAGGACCGCATGACG	721
Qy	434	SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly	453
Db	722	TCGTGTCATCGCATATGCTACAGAACCACCTCTCTGGCCTTTGTGGGCACCAAAAGTGGC	781
Qy	454	LysLeuLysLysIle	458
Db	782	AAGCTGAAGAAGGTG	796
RESULT 12			
US-10-238-370-91			
; Sequence 91, Application US/10238370			
; Publication No. US20030073191A1			
; GENERAL INFORMATION:			
; APPLICANT: Baker, Kevin			
; APPLICANT: Eaton, Dan			
; APPLICANT: Filvaroff, Ellen			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Grimaldi, J. Christopher			
; APPLICANT: Gurney, Austin			

Db	Seq ID	Accession	Organism	Length	Score	Similarity	Conservative
Db	542	TGTTACCGGGCGAGGCACGCTGGACCTGCCTGGCTCAAGGTGAAGGACATCCCTGTC	601	1.1e-138	2597	100.00%	1
QY	394	SerSerAlaLeuLeuThrIleAspAspAsnPhcCysGlyLeuAspMetAsnAlaProLeu	413	1.1e-138	1372.00	264	100.00%
Db	602	AGCAGTCCGCTCTTAACCAITGACGATAACTTCTGTGCCCTGGACATGAATGCTCCCTG	661	1.1e-138	1372.00	264	100.00%
QY	414	GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr	433	1.1e-138	1372.00	264	100.00%
Db	662	GGAGTGCCGACATGGTGGCGTGGAAATCCCGCTCTCACGGAGGACAGGACGCGATGAGC	721	1.1e-138	1372.00	264	100.00%
QY	434	SerValIleAlaIrrValTyLysAsnHisSerIleuAlaPheValGlyThrIlysserGly	453	1.1e-138	1372.00	264	100.00%
Db	722	TCTGTCATCGCATATGTCTACAAGAACCACCTCTCTGGCCTTTGTGGCCACCAAAAGTGC	781	1.1e-138	1372.00	264	100.00%
QY	454	LysLeuLysLysIle 458					
Db	782	AAGCTGAAGAAGGTG 796					
RESULT 11							
US-10-238-283-91							
; Sequence 91, Application US/10238283							
; Publication No. US20030073190A1							
; GENERAL INFORMATION:							
; APPLICANT: Baker, Kevin							
; APPLICANT: Eaton, Dan							
; APPLICANT: Filvaroff, Ellen							
; APPLICANT: Goddard, Audrey							
; APPLICANT: Grimaldi, J. Christopher							
; APPLICANT: Gurney, Austin							
; APPLICANT: Smith, Victoria							
; APPLICANT: Stephan, Jean-Phillippe							
; APPLICANT: Watanabe, Colin							
; APPLICANT: Wood, William							
; APPLICANT: Zhang, Zemin							
; APPLICANT: Fong, Sherman							
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC							
; TITLE OF INVENTION: ACIDS ENCODING THE SAME							
; FILE REFERENCE: P3630RJC15							
; CURRENT APPLICATION NUMBER: US/10/238, 283							
; CURRENT FILING DATE: 2002-09-09							
; PRIOR APPLICATION NUMBER: 10/197942							
; PRIOR FILING DATE: 2002-07-18							
; PRIOR APPLICATION NUMBER: 60/059114							
; PRIOR FILING DATE: 1997-09-17							
; PRIOR APPLICATION NUMBER: 60/063046							
; PRIOR FILING DATE: 1997-10-24							
; PRIOR APPLICATION NUMBER: 60/065027							
; PRIOR FILING DATE: 1997-11-10							
; PRIOR APPLICATION NUMBER: 60/079689							
; PRIOR FILING DATE: 1998-03-27							
; PRIOR APPLICATION NUMBER: 60/086478							
; PRIOR FILING DATE: 1998-05-22							
; PRIOR APPLICATION NUMBER: 60/087607							

```

; APPLICANT: Smith,Victoria
; APPLICANT: Stephan,Jean-Phillippe
; APPLICANT: Watanbe,Colin
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; APPLICANT: Fong,Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C10
; CURRENT APPLICATION NUMBER: US/10/238,370
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 91
; LENGTH: 2597
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-238-370-91

Alignment Scores:
Pred. No.: 1,1e-138 Length: 2597
Score: 1372.00 Matches: 264
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 0
Query Match: 13.73% Indels: 0
DB: 9 Gaps: 0

US-09-964-956-13 (1-1896) x US-10-238-370-91 (1-2597)
QY 194 GluTyrPheProThrIleSerArgLysLeuThrLysAsnSerGluAlaAspGlyMet 213
Db 2 GAGTATTTCCACCACATCTCCAGCGGAAACTGACCAAGAACTCTCAGCGGATGCGCATG 61
QY 214 PheAlaTyrValPheHisaspGluPheValAlaSerMetIleLysIleProSerAspThr 233
Db 62 TTCGGGTAGCTCTTCATGATGAGTTCGGGCTCGATGATTAAGATCCCTTCGGACACC 121
QY 234 PheThrIleIleProAspPheAspIleTyrTyrValTyrGlyPheSerSerGlyAsnPhe 253
Db 122 TTCACCAATCATCCCTGACCTTCTATCTACTATGCTATGGTTTATGACAGTGGCACTTT 181
QY 254 ValTyrPheLeuThrLeuGlnProGluMetValSerProGlySerThrThrLysGlu 273
Db 182 GTCTACTTTTGACCTCCCACTGAGATGGTGTCTCCACCGAGCTCCACCAAGGAG 241
QY 274 GlnValTyrThrSerLysLeuValArgLeuCysLysGluAspThrAlaPheAsnSerTyr 293
Db 242 CAGGTGTATACATCCAAAGCTCGTGGCTTTCGACGAGGACACAGCCTTCACTCCTAT 301
QY 294 ValGluValProIleGlyCysGluArgSerGlyValGluTyrArgLeuLeuGlnAlaAla 313
Db 302 GTAGAGGTGCCCATTTGGCTGTGAGCCAGTGGGGTGGAGTACCGCTCTGTCAGGCTGCC 361

314 TyrLeuSerLysAlaGlyAlaValLeuGlyArgThrLeuGlyValHisProAspAsp 333
Db 362 TACCTGTCCAAAGCGGGCGCTGCTTGGCAGGACCCTTGGAGTCCATCCAGATGATGAC 421
QY 334 LeuLeuPheThrValPheSerLysGlyGlnLysArgLysMetLysSerLeuAspGluSer 353
Db 422 CTGCTCTTCCACCGTCTTCTCCAAAGGCGCAGAGCGGAAATGAAATCCCTGGATGAGTCG 481
QY 354 AlaLeuCystIlePheIleLeuLysGlnIleAsnAspArgIleLysGluArgLeuGlnSer 373
Db 482 GCCTGTGCATCTTCATCTTGAAGCAGATAAATGACCGCATTAAGGAGCGCTGCAGTCT 541
QY 374 CysTyrArgGlyGluGlyThrLeuAspLeuAlaTrpLeuLysValLysAspIleProCys 393
Db 542 TGTACCGGGCGAGGCGACGCTGCACCTGGCTTCAAGGTCAAGGTGAAGGACATCCCTGTC 601
QY 394 SerSerAlaLeuLeuThrIleAspAspAsnPheCysGlyLeuAspMetAsnAlaProLeu 413
Db 602 AGCAGTGCCTCTTAACCATTTGACGATAAATCTGTGGCCTGGACATGAATGCTCCCTG 661
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspArgMetThr 433
Db 562 GGAGTGTCCGACATGGTGGTGGATTTCCCTCTTCACGGAGGACAGGACCGCATGACG 721
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValGlyThrLysSerGly 453
Db 722 TCTGTATCATGATATGCTACAGAACCACTCTCTGGCTTTGTGGGCGCCAAAGTGGC 781
QY 454 LysLeuLysLysIle 458
Db 782 AAGCTGAAGAAGGTG 796

RESULT 13
US-10-245-055-91
; Sequence 91, Application US/10245055
; Publication No. US20030073192A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanbe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C88
; CURRENT APPLICATION NUMBER: US/10/245,055
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
```





Db 602 AGCAGTGGCTCTTAACCATTTGACGATAACTTCTGTGGCTGGACATGAATGCTCCCTG 661  
QY 414 GlyValSerAspMetValArgGlyIleProValPheThrGluAspArgAspMetThr 433  
Db 662 GGAGTGTCCGACATGGTGGTGAATCCCGTCTTCACGGAGGACAGGACCGCATGACG 721  
QY 434 SerValIleAlaTyrValTyrLysAsnHisSerLeuAlaPheValClyThrLysSerGly 453  
Db 722 TCTGTATCGCATATGTCTACAAGAACCACTCTGTGGCCTTTGTGGGCACCAAAAGTGGC 781  
QY 454 LysLeuLysLysIle 458  
Db 782 AACCTGAAGAGGTG 796

Search completed: June 28, 2003, 19:18:57  
Job time : 749 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 28, 2003, 04:29:14 ; Search time 5336 seconds  
(without alignments)  
5754.618 Million cell updates/sec

Title: US-09-964-956-13  
Perfect score: 9990  
Sequence: 1 MKAMPNWTCLLSHLLVMG.....QKLYAKLEQVITLMSLDSNK 1896

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cnp2\_1/USPTO.spool/US09964956/runat\_24062003\_124339\_5214/app\_query.fasta\_1.2055  
-DB=EST -QPMF=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMF=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09964956.ecgn\_1.1.4160.@runat\_24062003\_124339\_5214 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_esthum:\*  
4: em\_esthum:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estro:\*  
9: gb\_esti:\*  
10: gb\_esti:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1843.5	18.5	4120	11	BC036274 Mus muscu
2	1390.5	13.9	2779	11	BC026974 Mus muscu
3	1357.5	13.6	1059	13	BM547417 AGENCOURT
4	1338	13.4	1027	13	BM544169 AGENCOURT
5	1299	13.0	980	9	AL529477 AL529477
6	1273	12.7	2691	11	BC011848 Homo sapi
7	1126	11.3	878	14	BQ732661 AGENCOURT
8	1119	11.2	627	12	BG086250 H3123F01-
9	1118	11.2	971	13	BI905192 603167521
10	1108	11.1	918	14	BQ723197 AGENCOURT
11	1096	11.0	844	14	BQ434703 AGENCOURT
12	1080	10.8	851	14	BQ878090 AGENCOURT
13	1077	10.8	762	9	AL134739 DXF2P547C
14	1072	10.7	742	12	BG294786 602391607
15	1066	10.7	612	10	BB656076 BB656076
16	1065.5	10.7	872	13	BI150919 602916727
17	1064.5	10.7	951	14	BQ920360 AGENCOURT
18	1064	10.7	920	14	BQ684665 AGENCOURT
19	1052.5	10.6	1174	13	BM479213 AGENCOURT
20	1052.5	10.5	933	14	BQ683335 AGENCOURT
21	1040.5	10.4	949	14	BQ678617 AGENCOURT
22	1039	10.4	725	14	BM950962 UI-M-EGO-
23	1038	10.4	961	14	BQ708128 AGENCOURT
24	1037	10.4	660	13	BM011540 603635809
25	1023	10.2	729	13	BQ072104 BJ072104
26	1022.5	10.2	881	14	BQ680140 AGENCOURT
27	1021.5	10.2	664	14	BQ180472 UI-M-EXO-
28	1019.5	10.2	906	14	BQ947921 AGENCOURT
29	1005	10.1	904	14	BQ720660 AGENCOURT
30	994.5	10.0	731	14	BQ444249 UI-M-EXO-
31	994.5	10.0	838	13	BI824913 603032430
32	990	9.9	999	14	BQ073659 AGENCOURT
33	981.5	9.8	702	12	BG825770 602747116
34	975	9.8	957	12	BF308474 601887611
35	974.5	9.8	927	14	BQ897297 AGENCOURT
36	973.5	9.7	924	12	BF312056 601897930
37	973.5	9.7	994	12	BE736103 601305855
38	970	9.7	653	10	BE294546 601173485
39	965	9.7	659	14	BQ394478 NISC_ng09
40	960	9.6	748	12	BF311719 601897214
41	959.5	9.6	662	14	BQ180469 UI-M-EXO-
42	947	9.5	732	10	BE269288 601186342
43	945	9.5	634	14	BM936540 UI-M-BH3-
44	940	9.4	710	10	BE295416 601175942
45	934.5	9.4	951	12	BF309302 601887769

ALIGNMENTS

RESULT 1  
BC036274  
LOCUS BC036274  
DEFINITION Mus musculus, clone IMAGE:3588465, mRNA.  
ACCESSION BC036274  
VERSION BC036274.1  
KEYWORDS HTCC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 4120)  
Strausberg,R.  
Direct Submission  
TITLE Submitted (01-AUG-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAP Plate: 10 Row: J Column: 8

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis

This clone has the following problem: no 5' EST match.

#### FEATURES

##### source

1. 4120

Location/Qualifiers

/organism="Mus musculus"

/db\_xref="taxon:10090"

/map="CZCH 11"

/clone="IMAGE:3588465"

/tissue\_type="Mammary tumor metastasized to lung. Tumor

arose spontaneously from a senescent normal mammary

(clonal) outgrowth infected with the virus MMTV."

/clone\_lib="NCI CGAP\_Lu29"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

953 a 1166 c 1151 g 850 t

#### BASE COORDINATE

##### ORIGIN

#### Alignment Scores:

Pred. No.: 2,076-210 Length: 4120  
Score: 1843.50 Matches: 444  
Percent Similarity: 54.36% Conservative: 204  
Best Local Similarity: 37.25% Mismatches: 403  
Query Match: 18.45% Indels: 141  
DB: 11 Gaps: 33

US-09-964-956-13 (1-1896) x BC036274 (1-4120)

QY	737	SerSerValGlnCysGlnAsnThrSerTyrSerTyrGluGlyMetGluLeuAsnAsn	776
DB	2	TCAGACACCTCTCTTTTAGGACCCCAAGCTATCTATGATGTTAATGAG-----ACA	55
QY	777	LeuProValGluLeuThrValValTrpAsnGlyHisPheAsnIleAspAsnProAlaGln	796
DB	56	CTGCCTCTTACCTCTATGTTAAGTCCTTTGGCAAG---AACATTGACAGCAAGCTACAA	112
QY	797	AsnLysValHisLeuTyrLysCysGlyAlaMetArgGluSerCysGlyLeuLys	816
DB	113	-----GTGACTCTCTACACTGCTCTTGGCCGAGTGACTGTAGCTGTCTGGCT	166
QY	817	AlaAspProAspPheAlaCysGlyTrpCysGlnGlyProGlyGlnCysThrLeuArgGln	836
DB	167	GCGGATCTGCTACAGGTGTGTGGTGGCGGAGAACCGGTGCGGTGACGAAGCT	226
QY	837	HisCysProAlaGlnGluSerGlnTrpLeuGluLeuSerGlyAlaLysSerLysCysThr	856
DB	227	CTGTGC-----AGCAATGCTCACTCTGAGTGCCCA	256
QY	857	AsnProArgIleThrGluIlelleProValThrGlyProArgGluGlyThrLysVal	876
DB	257	CCACCAGTATACATAGGATCCAGCCTGAGAGCGGCCCGCTGGGTGGGGGCATCCTGTC	316

QY	877	ThrIleArgGlyGluAsnLeuGlyLeuGluPheArgAspIleAlaSerHisValVal	896
DB	317	ACTATCCATGGTCCCAATCTGGTGTCAAAGCAGATGACGTC---AAGAAGATAACTGTG	373
QY	897	AlaGlyValGluCysSerProLeuValAspGlyTyrIleProAlaGluGlnIleValCys	916
DB	374	GCTGCCAGAACTGTGCTTTGAACCAAGAGGATCTCGTATCCACCCCGATTTGTGT	433
QY	917	GluMetGlyGluAlaLysProSerGlnHisAlaGlyPheValGluIle--CysValAla	936
DB	434	GCAATTT---GAGCTTCGAGATGCCCTTCACAGGAGGATTCAGGTGGATGTTAATGGA	490
QY	936	aIcysArgProGluPheMetAlaArgSerGlnLeuTyrTyrPheMetThrLeuThrL	956
DB	491	AAACTCGGCCA---TTCACCGGCCAC---GTCAGTTTCACTTATACACACCCAGCCTC	544
QY	956	eUeSerAspLeuLysProSerArgGlyProMetSerGlyGlyThrGlnValThrIleThr	976
DB	545	TCAGT---GTGGAGCCACACAGGGGCCACAGGAGGTGGCCACACATGACCATCATG	601
QY	976	lyThrAsnLeuAsnAlaGlySer-----AsnValValValMetPheGlyLysGlnProc	994
DB	602	GCACTCACCTGGACACAGGCTCCCAAGGAGGATGCGGGTGACACTCAATGACGCTCCT	661
QY	994	ysLeuPheHisArgArgSerProSerTyrIleValCysAsnThrThrSerSerAspGlu	1014
DB	662	GTGAGTG---ACAAAGTTTGGAGCACAGCTACAGTGTGCACAGGTCAACAGTTGGCTC	718
QY	1014	alleuGluMetLysValSerValGlnValAspArgAlaLysIle---HisGlnAspLeu	1033
DB	719	CAGGCCAGGTGACACTAGAAATCTACTATGGAGTCCACAGTCCCGCCGATCT	778
QY	1033	alPheGlnTyrValGluAspProThrIleValArgIleGluProGluTrpSerIleVal	1053
DB	779	CITTCACCTACTCGAGAACCCCATGATACGAGCTTTTGAGCCATTGAGAAGCTTTGTCA	838
QY	1053	cTtGlyAsnThrProIleAlaValTrpGlyThrHisLeuAspLeuIleGlnAsn-----	1070
DB	839	GTGGTGGCCGAGCATCACGTTACTGGCCAGGCTTCAGCCTCATCCAGAAGTTTGCCA	898
QY	1071	-----ProGlnIleArgAlaLysHisGlyGlyL	1080
DB	899	TGTTTTCATCGTACCCCTTGGGTCCTGGAGCGCGCGGAGGCTGGAGGCC	958
QY	1080	ysGluHisIleAsnIle-----CysGluValLeuAsnAlaThrGluMetThrC	1096
DB	959	TGGAGCGGTGTGACGGTCGAGGCGATGGAGTACGTGTTCTACATGACACCAAGGTCGT	1018
QY	1096	ysGlnAlaProAlaLeuAlaLeuGlyProAspHisGlnSerAspLeuThrGluArgPro	1116
DB	1019	TCTTGCTCTCTGCT-----GTCCCCAAGAGGCTG	1048
QY	1116	luGluPheGlyPheIle-----LeuAspAsnValGlnSerLeuLeuLeuA	1132
DB	1049	AGGCTTACACCTCACTGTCTGATACGGATGAGCGGTCACTGTGCCCGCTCAGGACCG	1108
QY	1132	snLysThrAsnPheThrTyrProAsnProValPheGluAlaPheGlyProSerGlyI	1152
DB	1109	AAGCCGGAGTCTTTGAGTATGTGCAGACCCCATCTTGAAGACTTC-----	1155
QY	1152	leLeuGluLeuLysProGlyThrProIleLeuLysGlyLysAsnLeuIleProPro	1172
DB	1156	-----ACGGTGGCGTCAAAAGCAGGTCAACAGCTCATCCACG	1195
QY	1172	alAlaGlyGlyAsnValLysLeuAsnTyrThrVal-----LeuValGlyG	1187
DB	1196	CCCGGGAACCAATCTGAACAAAGCATGACTCTGGAGGAGGCGAGAGGCTTTGGGTG	1255
QY	1187	lulysProCysThrVal---ThrValSerAspValGlnLeuLysGluSerProAsnL	1206
DB	1256	CTGAGCGTTCATTTATGAAGACGCTGACCGAGAGGACCTGTACTGCGAGGCCCGCAGG	1315



LOCUS BC026974 2779 bp mRNA linear HTC 07-AUG-2002  
DEFINITION Mus musculus, clone IMAGE:5364076, mRNA.  
ACCESSION BC026974  
VERSION BC026974.1 GI:20072031  
KEYWORDS HTC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 2779)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 53 Row: g Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis  
This clone has the following problem: incomplete processing.  
FEATURES  
source  
1..2779  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5364076"  
/tissue\_type="Eye, retina, mouse strain C57Bl/6"  
/clone\_lib="NIH\_MGC\_94"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 700 a 717 c 766 g 596 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.63e-156 Length: 2779  
Score: 1390.50 Matches: 273  
Percent Similarity: 66.61% Conservative: 104  
Best Local Similarity: 48.23% Mismatches: 146  
Query Match: 13.92% Indels: 43  
DB: 11 Gaps: 8  
US-09-964-956-13 (1-1896) x BC026974 (1-2779)  
QY 1336 ArgAspLeuValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPhe 1355  
Db 5 AGGGACCTCGGTGCTGCACACAGCGGCCACCGTGGAAACAGGCGCTGGGGCAGCTC 64  
QY 1356 AlaGlnLeuLeuAsnLysValPheLeuLeuSerPheLeuArgThrLeuGluSerGln 1375  
Db 65 TCCAACTGCTAAACAGCAGCTCTTCCCTACCAAGTTTCATCCACACACTGGAGAGTCAG 124  
QY 1376 ArgSerPheSerMetArgAspArgGlyAsnValAlaSerLeuLeuMetThrValLeuGln 1395  
Db 125 CGACCTTCCTCTCGTGGGACCGGCGCTACGTGGCATCTCTGTCACCTGTGCACCTCAC 184  
QY 1396 SerLysLeuGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuLys 1415

185 GGGAGCTGTAATACTTACGACACATACTGGGACTCTGCTCAGTGCACCTGTAGCTCAG. 244  
Db  
1416 AsnLeuGluSerLysAsnHisProLysLeuLeuArgArgThrGluSerValAlaGlu 1435  
QY  
245 TATGTT---GCCAAGAAC---CCCAAGCTGATGTCGCAGACAGACAGACCGTGGTAGAA 298  
Db  
1436 LysMetLeuThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGly 1455  
QY  
299 AAGCTGCTCAACCACTGGATGCCCTCTACACCTTTGTGAGGAGCTGTGTGGGA 358  
Db  
1456 GluProLeuPheSerLeuPheCysAlaLeuLysGlnMetGluLysGlyProLeuAsp 1475  
QY  
359 GAGCCTCTGTATATGCTCTTCAGAGGGATTAAAGCATCAAGTGGACAAGGTCCCGTGGAC 418  
Db  
1476 AlaLeuThrGlyGluAlaArgTyrSerLeuSerGluAspLysLeuLeuArgGlnGlnIle 1495  
QY  
419 AGTGTGACTGGCAAGCCCAATACACTCTGAATGACACACCGCTGCTCAGAGGAGTGTG 478  
Db  
1496 AspTyrLysThrLeuValLeuSer-----CysValSerProAspAsnAlaAsnSer 1512  
QY  
479 GAGTACCGTCCTTGCACCTTGAATGCTCTTGGCTGTGGGCGCTGGGCGAGGAGGCC 538  
Db  
1513 ProGluValProValLysLeuLeuAsnCysAspThrIleThrGlnValLysGluLysIle 1532  
QY  
539 CAGTGTGTACTGTGAAAGTCTCTGGACTGTGACACCATCTCCAGGCCAAGGAGAGATG 598  
Db  
1533 LeuAspAlaIlePheLysAsnValProCysSerHisArgProLysAlaAlaAspMetAsp 1552  
QY  
599 CTAGACCACTTTACAAGGAGTGCCTCTTGGCCAGCGCGCCGACTCTGCACCTTGGAT 658  
Db  
1553 LeuGluTrpArgGlnGlySerGlyAlaArgMetIleLeuGlnAspIleThrThr 1572  
QY  
659 GTTGAATGGCGGCTCTGGAGTGGCGCCACTTATCTCTTCTGATGAGGAGCTCCTCC 718  
Db  
1573 LysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHisTyrGlnValProAspGly 1592  
QY  
719 GAATCCAGGCTCTGTGGAGCGCTGATACACTGCAACATTAACAAGTCCCATGGA 778  
Db  
1593 SerValValAlaLeuVal-----SerLysGlnValThrAlaTyrAsnAlaValAsn 1609  
QY  
779 GCAACGCTGGCCCTTCTCCCTGCTCCCAAGCATATTCTTAGG----- 823  
Db  
1610 AsnSerThrValSerArgThrSerAlaSerLysTyrGluAsnMetIleArgTyrThrGly 1629  
QY  
824 -----GAAACACCAAGATTATGTCCTCCTGGG 847  
Db  
1630 SerProAspSerLeuArgSerArgThrProMetIleThrProAspLeuGluSerGlyVal 1649  
QY  
848 -----GAACGGACCCCAATGCTGGAGGATGTAGATGAGGGGGGCATC 889  
Db  
1650 LysMetTrpHisLeuValLysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGly 1669  
QY  
890 CGGCCCTGGCACCTGGTAAAGCCAGCATGATGAACAGAGCCCTCCAGGCCGAGGAGGCG 949  
Db  
1670 Ser-----LysMetValSerGluIleTyrLeuThrArg 1680  
QY  
950 AGCCTTCGGGTGGGAGCGGTGAGCGAGCCCAAGGCTATCCCTGAGATCTACCTGACACGC 1009  
Db  
1681 LeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspLeuPheGluThrIlePhe 1700  
QY  
1010 CTGTATCCATGAAGGGCACACTGCAAGAGTTGTGGATGACCTGTTCCAGGTGATCTC 1069  
Db  
1701 SerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeu 1720  
QY  
1070 AGCACCAGCCG-----CCTGTGCTCTGGCTGTGAGTACTTCTTGTGACTGTGCTG 1120  
Db  
1721 AspGluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrpLysSer 1740  
QY  
1121 GATGAACAAGCTCAGCAGCATGGCATCTCTGTATCAGGATACACTATCCATCTGGAAGACC 1180  
Db  
1741 AsnCysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheValPheAsp 1760  
QY

Db 1181 AACAGCTCGCGCTAGGTTCTGGATCAACATCATCAAGAACCCACAGTTTGTGTCGAT 1240  
QY 1761 IleHisLysAsnSerIleThrAspAlaCysLeuSerValAlaGlnThrPheMetAsp 1780  
Db 1241 GTGCAGACTTCGATAAATGATGATGCTGCTCTGTCATTCACAGACCTTCATGGAT 1300  
QY 1781 SerCysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyr 1800  
Db 1301 GCTTGACCTCGCGCAGCACCAGCTGCGCGGATTCCTCCCATCAACAACCTCTGTAT 1360  
QY 1801 AlaLysAspIleProSerTyrLysAsnTyrValGluArgTyrSerAspIleGlyLys 1820  
Db 1361 GCTCGAGATATCCCGTTCACAAAGAGTGGTGAAGGTAATGACGACATCAGACAG 1420  
QY 1821 MetProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGlnSerArgMetHis 1840  
Db 1421 ACTGTCCCGCAGTGACCAAGAGATGAATCACTGTGCGGAGCTGTCCCGGAAGTGC 1480  
QY 1841 MetAsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGlyLysTyr 1860  
Db 1481 TCTGCGACCTTGGGCGCGAGTGGCTCTGCATGAATCTACAAAGTATATCAACAAGTAC 1540  
QY 1861 SerGluGluIleLeuGlyProLeuAspHisAspGlnCysGlyLysGlnLysLeuAla 1880  
Db 1541 TATGACCATATCATCACTCCCTGGAGAGGATGGCACTGCCAGAGATGCGAGCTGGGC 1600  
QY 1881 TyrLysLeuGluGlnVal 1886  
Db 1601 TACCGGCTCCAGCAGATC 1618

RESULT 3  
BM547417

LOCUS BM547417 1059 bp mRNA linear EST 20-FEB-2002  
DEFINITION AGENCOURT\_6507625 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5724374  
5' mRNA sequence.

ACCESSION BM547417  
VERSION BM547417.1 GI:187811167  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1059)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12713 row: i column: 15  
High quality sequence stop: 713.  
Location/Qualifiers  
1. .1059  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5724374"  
/lab\_host="NIH\_MGC\_125"  
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site1: EcoRV (destroyed); Site2: NotI; RNA source pool  
of three ovaries, from females ranging in age from 38 to  
49 yo. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 2.1 kb, insert size range 1-3.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 036."

FEATURES  
source

BASE COUNT 250 a 308 c 303 g 198 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,33e-152 Length: 1059  
Score: 1357.50 Matches: 278  
Percent Similarity: 89.29% Conservative: 22  
Best Local Similarity: 82.74% Mismatches: 30  
Query Match: 13.59% Indels: 6  
DB: 13 Gaps: 2  
US-09-964-956-13 (1-1896) x BM547417 (1-1059)  
QY 1256 ValLeuIleAlaTyrLysArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGln 1275  
Db 15 GTCCTCATTTGCTACAAAGCGAAGTCTCGAAGAAATGACCTCACCTCTCAAGCGGTGCAA 74  
QY 1276 MetGlnMetAspAsnLeuGluSerArgValAlaLeuGluCysLysGluAlaPheAlaGlu 1295  
Db 75 ATGCAGATGGACAATCTGGAGTCCCGTGTGGCTTGGAGTGCAGGAAGCTTTTGTGTGAG 134  
QY 1296 LeuGlnThrAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeu 1315  
Db 135 CTCCAGAGGATATCAATGAGTTGACCACTGACCTGGACCGCTCAGGAATCTCTTACCTG 194  
QY 1316 AspTyrArgThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLeu 1335  
Db 195 GACTATCGTACCTACGCTATCGAGTCTCTGTCGGGCGCATCGAGGACCAACCCCGCTCTG 254  
QY 1336 ArgAspLeuGluValProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPhe 1355  
Db 255 CGGAGCTGGAGGTACAAAGAACGGGACGACGACGTGGAGAAAGGCGCTGAAGCTCTTT 314  
QY 1356 AlaGlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGln 1375  
Db 315 GCCCAGCTCATCAACACAAAGGTGTCTGTGACCTTCATCCGACACCTCGAGCTGCAG 374  
QY 1376 ArgSerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGln 1395  
Db 375 CGCAGTTTCTCCATGCGCGACCGGCGAACGTGGCTTCGCTCATCATGACCGCGCTGCAG 434  
QY 1396 SerLysLeuGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuLysLys 1415  
Db 435 GGCCGCTTGGAAATATGCCACTGATGCTCTCAAGCAGCTGCTCTGACCTCATCGATAAG 494  
QY 1416 AsnLeuGluSerLysAsnHisProLysLeuLeuLeuArgArgThrGluSerValAlaGlu 1435  
Db 495 AACCTGGAGACACAGAACCCACCCCAAGCTGCTACTCCGAGGACAGAGTCTGTGGCTGAA 554  
QY 1436 LysMetLeuThrAsnTyrPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGly 1455  
Db 555 AAGATGCTGACCAATTTGGTTTCGCTTCCTCTCGACAAAGTTCTTAAGAGAGTGCAGGG 614  
QY 1456 GluProLeuPheSerLeuPheCysAlaIleLysGlnMetGluLysGlyProIleAsp 1475  
Db 615 GAGCCACTCTTTCATGCTATATGTCATCAAGCAGCAGATGGAGAGGGGCCCATTTGAT 674  
QY 1476 AlaIleThrGlyLysAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnIle 1495  
Db 675 GCCATCAGGCGAGGCGCGCTACTCTGAGGAGGACAAAGCTCATCGCGCAGCAGATC 734  
QY 1496 AspTyrLysThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluVal 1515  
Db 735 GAGTACAAGACCTTGATCTCTGAACCTGCTCAACCTGACACGACGACAGTCCACGAGATC 794  
QY 1516 ProValLysIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAla 1535  
Db 795 CCCAGTGAAGTGTAAACTGTGACACCATCCCGCAGGTCGAGGAGAAAGATTCTTGTATGCC 854  
QY 1536 IlePheLysAsnVal-ProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTr 1555  
Db 855 GGGGATAGAATGTGGCCCTATTCCACGCGCGGAGGCGAGTGGACATGGACTTGGGATG 914

QY	1555	pArgGInGlySerGly-AlaArgMetIle-LeuGlnAspGluAspIleThrThrLysIle	1574
Db	915	GCAGCAAGGCGCGGATCGCGGCTCGGGCTTGAGATGAGGACATCCCGCCAGATT	974
QY	1575	GluAsnAspTrp-----LysArgLeuAsnThrLeuAlaHis	1586
Db	975	GAGAGG---TGGACTGGGAACGCTCAACACCTGAGTGAT	1013
RESULT 4			
BM544169			
LOCUS	BM544169	1027 bp mRNA linear	EST 20-FEB-2002
DEFINITION	AGENCOURT_6490555 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5587657		
ACCESSION	BM544169		
VERSION	BM544169.1	GI:18775199	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	NIH (bases 1 to 1027)		
TITLE	NH-MGC http://mgi.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs@mail.nih.gov		
	Tissue Procurement: Invitrogen		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: L1AM12357 row: 1 column: 02		
	High quality sequence stop: 676.		
FEATURES			
Source	1..1027		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5587657"		
	/clone_lib="NIH_MGC_125"		
	/lab_host="DH10B"		
	/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;		
	Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool		
	of three ovaries, from females ranging in age from 38 to		
	49 yo. Library is oligo-dT primed and directionally cloned		
	(EcoRV site is destroyed upon cloning). Average insert		
	size 2.1 kb, insert size range 1-3.5 kb. Library is		
	normalized and enriched for full-length clones and was		
	constructed by C. Gruber (Invitrogen). Research Genetics		
	tracking code 036."		
BASE COUNT	238 a 317 c 303 g 169 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.9e-150	Length:	1027
Score:	1338.00	Matches:	270
Percent Similarity:	87.02%	Conservative:	25
Best Local Similarity:	79.65%	Mismatches:	38
Query Match:	13.39%	Indels:	7
DB:	13	Gaps:	2
US-09-964-956-13 (1-1896) x BM544169 (1-1027)			
QY	1430	ThrGluSerValAlaGluLysMetLeuThrAsnTrpPheThrPheLeuLeuTrLysPhe	1449
Db	20	TCAGTCGGTGGCAGAGATGCTAACTAAGTTCACCTTCCTCTGTATAAGTTC	79
QY	1450	LeuLysGluCysAlaGlyGluProLeuPheSerLeuPheCysAlaIleLysGlnMet	1469
Db	80	CTCAAGGAGTGCCTGGGAGCGGCTGTTTCATCTGCTACTGCGCCATCAAGCAGCATG	139
QY	1470	GluLysGlyProIleAspAlaIleThrGlyGluAlaArgTyrSerLeuSerGluAspLys	1489

TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

## source

Location/Qualifiers  
1..980  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSDD006YE12"  
/clone\_lib="LRI\_NFL001\_NBC4"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 241 a 276 c 272 g 188 t 3 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1,43e-145 Length: 980  
Score: 1299.00 Matches: 271  
Percent Similarity: 90.80% Conservative: 25  
Best Local Similarity: 83.13% Mismatches: 26  
Query Match: 13.00% Indels: 8  
DB: 9 Gaps: 1

US-09-964-956-13 (1-1896) x AL529477 (1-980)

QY 1261 LysArgLysSerArgGluSerAspLeuThrLeuLysArgLeuGlnMetGlnMetAspAsn 1280  
DB 1 AAGCCAAAGTCTCGAGAAATGACCTCACTCTCAAGCGGCTCAAAATGCAGATGGACAT 60  
QY 1281 LeuGluSerArgValAlaLeuGluCysLysGluAlaPheAlaGluLeuGlnThrAspIle 1300  
DB 61 CTGGAGTCCGCTGGCCCTGGAGTSCAAGGAAGCTTTTCTGAGCTCCAGACGGATATC 120  
QY 1301 HisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeuAspThrArgThrTyr 1320  
DB 121 AATGATGTACAGTACCTGGACCGCTCAGGAATCCCTTACTGGACTATCGTACCTAC 180  
QY 1321 ThrMetArgValLeuPheProGlyIleGluAspHisProValLeuArgAspLeuGluVal 1340  
DB 181 GCTATGCGAGTCTCTTCCCGGCATCGAGGACCAACCCGCTCTCGGGGAGCTGGAGGTA 240  
QY 1341 ProGlyTyrArgGlnGluArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIleAsn 1360  
DB 241 CAAGGAACCGGCACACACCTGGAGAGGCCCTTGAAGCTCTTTGCCAGCTCATCAAC 300  
QY 1361 AsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGlnArgSerPheSerMet 1380  
DB 301 AACAGGTGTCTCTGCTGACCTTCATCCGACCCCTGGAGCTCGAGCGAGTTTCTCCATG 360  
QY 1381 ArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGlnSerLysLeuGluTyr 1400  
DB 361 CGCGACCGGGCAACGTGGCTTCGCTCATCATGACCGGCTCCAGGGCSC-CTGGAATAT 419  
QY 1401 AlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLysAsnLeuGluSerLys 1420  
DB 420 GCCACTGATGTCTCAAGCAGCTGCTCTCTGACCTCATCGATAAAGAACCTGGAGAACAG 479  
QY 1421 AsnHisProLysLeuLeuLeuArgThrGluSerValAlaGluLysMetLeuThrAsn 1440  
DB 480 AACCAACCCCAAGCTCTACTCGGAGGACAGAGCTGTGGCTTGAAAGATGCTGACCAAT 539

QY 1441 TrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGlyGluProLeuPheSer 1460  
DB 540 TGGTTCGCCTTCCTCTGTCACAAAGTCTCTAAAGAGAGTGGCAGGGAGCCACTCTTCATG 599  
QY 1461 LeuPheCysAlaIleLysGlnMetGluLysGlyProIleAspAlaIleThrGlyGlu 1480  
DB 600 CTATACTGTGCCATCAAGCAGCAGATGGAGAAGGGCCCATTTGATGCGATCACCAGCGAG 659  
QY 1481 AlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIleAspTyrLysThrLeu 1500  
DB 660 GC-CGCTACTCTCTGAGCAGGACAGCTATCCGCGAGCAGATCGAGTCAACAGCCCTG 718  
QY 1501 ValLeuSerCysValSerProAspAsnAlaAsnSerProGluValProValLysIleLeu 1520  
DB 719 ATCTCTGAACCTGCTCAACCTCTGACACGAGAGAACAGTCCAGAGATCCAGTGAAGTGTA 778  
QY 1521 AsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAlaIlePheLysAsnVal 1540  
DB 779 AACTGTGACACCATCAACAGGTCACAGGAGAAATCTTGTATGCGCTGTATGAAGAATGTG 838  
QY 1541 ProCysSerHisArgProLysAlaAlaAspMetAspLeuGluTyrArgGlnGlySerGly 1560  
DB 839 CCTATTCCAGCGCGCGAGGCGAGTGCACATGAGCTTGGAGTGGCG-CAAGGCCGGATC 897  
QY 1561 AlaArgMetIleLeuGlnAspGluAspIleThrThrLysIleGluAsnAspTrpLysArg 1580  
DB 898 GCCCGGCTGCTGTCARAT-GAGGACATCACCACAAGATTG-----AGGCTR 944  
QY 1581 LeuAsnThrLeuAlaHis 1586  
DB 945 CTGGAGCGGCTCACAC 962  
RESULT 6  
BC011848  
LOCUS Homo sapiens, clone IMAGE:4121608, mRNA.  
DEFINITION BC011848  
ACCESSION BC011848.1 GI:15341708  
VERSION HTG  
KEYWORDS HTG  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2691)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland:  
Web site: http://www.nisc.nih.gov/  
Contact: nisc\_mgc@hgrl.nih.gov  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Brooks, S.,  
Benjamin, B., Blakesley, R.W., Bouffard, G., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,  
Tsongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
Zhang, L.H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAL Plate: 28 Row: i Column: 23

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis  
This clone has the following problem: frame shifted.

FEATURES		Location/Qualifiers	
source	1..2691	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:4121608"	
		/tissue_type="Muscle, rhabdomyosarcoma"	
		/clone_lib="NIH_MGC_17"	
		/lab_host="DH10B-R"	
		/note="vector: pOTB7"	
BASE COUNT	632 a 776 c 777 g 506 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1..45e-141	Length:	2691
Score:	1273.00	Matches:	268
Percent Similarity:	65.30%	Conservative:	97
Best Local Similarity:	47.94%	Mismatches:	133
Query Match:	12.74%	Indels:	61
DB:	11	Gaps:	10
US-09-964-956-13 (1-1896) x BC011848 (1-2691)			
Qy	1347	ArgValGluLysGlyLeuLysLeuPheAlaGlnLeuIleAsnAsnLysValPheLeuLeu	1366
Db	5	CGAGGGGAAGAGGAATTAGCTTCTCCATCTACTCAACACAGACACTTCTCATC	64
Qy	1367	SerPheIleArgThrLeuGluSerGlnArgSerPheSerMetArgAspArgGlyAsnVal	1386
Db	65	GTCCTTGTCCACGCGCTGGACGACGAGAGGACCTTGGCGGTGCGCAGAGTGCAGCCTG	124
Qy	1387	AlaSerLeuIleMetThrValLeuGlnSerLysLeuGluTyrAlaThrAspValLeuLys	1406
Db	125	CCCTCGCTGTGACCATCGCGCTGCACGCGAGCTGGAGTACTACACGACATCAATGAG	184
Qy	1407	GlnLeuLeuAlaAspLeuIleAspLysAsnLeuGluSerLysAsnHisProLysLeuLeu	1426
Db	185	GAGCTGTGTTGGACCTCATTGACGCTCG-----GCCGCCAAGAACCCCAAGCTCATG	238
Qy	1427	LeuArgArgThrGluSerValAlaGluLysMetLeuThrAsnTrpPheThrPheLeuLeu	1446
Db	239	CTGCGCGCACAGAGTCTGTGTGGAGAGAGATGCTCACCAACTGGATGTCATCTGCATG	298
Qy	1447	TyrLysPheLeuLysGlyCysAlaGlyGluProLeuPheSerLeuPheCysAlaIleLys	1466
Db	299	TACAGCTGTCTCGGGGAGACGGTGGGGAGGCCATTCTTCTCTGCTGCTGTGCGCATCAG	358
Qy	1467	GlnGlnMetGluLysGlyProIleAspAlaIleThrGlyGluAlaArgTyrSerLeuSer	1486
Db	359	CAGCAATCAACAGGGCTCCATGACGCGCATCACAGGCAAGCGCGCTACACACTCAAT	418
Qy	1487	GluAspLysLeuIleArgGlnClnIleAspTyrLysThrLeuValLeuSerCysValSer	1506
Db	419	GAGGAGTGGCTGCTCGGGAGAACATCGAGGCGCAACCCCGCAACCTGAACTC---G	475
Qy	1507	ProAspAsnAlaAsnSerProGluValProValLysIleLeuAsnCysAspThrIleThr	1526
Db	476	TTCCAGGGCTGTGGCATGCTGAGCGGTGCGGGCCATGGACACGACGCTGACA	535
Qy	1527	GlnValLysGluLysIleLeuAspAlaIlePheLysAsnValProCysSerHisArgPro	1546
Db	536	CAGGTCAAGGAGAAGATCTCGAGGCTTCTCAAGAATGTGCCCTACTCCAGTGGCCG	595
Qy	1547	LysAlaAlaAspMetAspLeuGluTrpArgGlnGlySerGlyAlaAlaArgMetIleLeuGln	1566
Db	596	CTGTGCAGGAGAGTGCACCTTGAGTGTGCTCCGCTCCAGCACACAGAGCTACATCTCG	655
Qy	1567	AspGluAspIleThrThrLysIleGluAsnAspTrpLysArgLeuAsnThrLeuAlaHis	1586
Db	656	GACCTGGAGACACCTCAGTGTGTGGAAGAGCGCGCGCAAGAAGCTTAACAGCTGGCC	715

Qy	1587	TyrGlnValProAspGlySerValValAla-----LeuValSerLysGlnValThrAla	1604
Db	716	TACAAGATCCCTGAAGTGCCTCCCTGGCCATGACTCTCATAGACAAGAAG-----	766
Qy	1605	TyrAsnAlaValAsnAsnSerThrValSerArgThrSerAlaSerLysTyrGluAsnMet	1624
Db	767	-----GACAACACACTGGCGCGAGTGAAA-----	790
Qy	1625	IleArgTyrThrGlySerProAspSerLeuArgSerArgThrProMetIleThrProAsp	1644
Db	791	-----GAC	793
Qy	1645	LeuGluSerGlyValLysMetTrpHisLeuVal-----	1655
Db	794	TTGGACACA---GAGAAGTATTTCCATTTGGTGTGCTTACGACGACGAGTGGCGGAGCCC	850
Qy	1656	---LysAsnHisGluHisGlyAspGlnLysGluGlyAspArgGlySerLysMetValSer	1674
Db	851	AAGAAGTCTCACCGCGCAGAGCCATCGCAAG-----AAGGTGCTCCCG	892
Qy	1675	GluIleTyrLeuThrArgLeuLeuAlaThrLysGlyThrLeuGlnLysPheValAspAsp	1694
Db	893	GAATCTTACCTGACCGCGCTGCTCCACCAAGGCGACGTTCAGAAAGTTTCTGGATGAC	952
Qy	1695	LeuPheGluThrIlePheSerThrAlaHisArgGlySerAlaLeuProLeuAlaIleLys	1714
Db	953	CTGTTCAAGGCCATTTCTGATATC-----CGTGAAGACAAGCCCCACTGGGTGTCAAG	1006
Qy	1715	TyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIleHisAspProHisVal	1734
Db	1007	TACTTTTTCGACTTCTGGAGGAGCAGGTGAGAAAGGGGAATCTCCGACCCGACACC	1066
Qy	1735	ArgHisThrTrpLysSerAsn-----CysLeuProLeuArgPheTrpValAsnMetI	1752
Db	1067	CTACACATCTGGAAGACCAACAGGTGGCGGCTTCTCTCGGTCTTGGGTGAACATCCT	1126
Qy	1752	eLysAsnProGlnPheValPheAspIleHisLysAsnSerIleThrAspAlaCysLeuSe	1772
Db	1127	GAAGAACCCCGAGTTGTCTTTCATCATCGACAGACACACACATCGACCGCTGCTTC	1186
Qy	1772	rValValAlaGlnThrPheMetAspSerCysSerThrSerGluHisArgLeuGlyLysAs	1792
Db	1187	AGTCATCGCGCAGCGCTTCATCGACGCTGCTCCATCTCTGACCTGAGCTGGCAAGGA	1246
Qy	1792	pSerProSerAsnLysLeuLeuTyrAlaLysAspIleProSerTyrLysAsnTrpValG	1812
Db	1247	TTGCGCAACCAACAAGCTCTCTACGCCAAGGAGATTCTCTAGTACCGAAGATCGTGCA	1306
Qy	1812	uArgTyrTyrSerAspIleGlyLysMetProAlaIleSerAspGlnAspMetAsnAlaTy	1832
Db	1307	GCCTACTACAACAGCATCCAGGCATGACCGCGCTCAGCGAGCAAGAGATGAATGCCA	1366
Qy	1832	rLeuAlaGluGlnSerArgMetHisMetAsnGluPheAsnThrMetSerAlaLeuSerG	1852
Db	1367	TCTGGCGGAGGAGTCGAGGAATACCAAGATGAGTTCAACACCAATGTGGCCATGGCAG	1426
Qy	1852	uIlePheSerTyrValGlyLysTyrSerGluGluIleLeuGlyProLeuAspHisAspAs	1872
Db	1427	GATTTATAAGTACCCCAAGAGTATCGCGCCGACATCATGCCCGCTGGAGGCCAACCC	1486
Qy	1872	pGlnCysGlyLysGlnLysLeuAlaTyrLysLeuGluGlnValIleThrLeuMet	1890
Db	1487	CACGGCCCGGAGGACACAACACTGCAGCACACAAGTTTGAGCAGGTGGTGGTTTGTATG	1541
RESULT 7			
BQ732661			
LOCUS			
DEFINITION			
BQ732661			
ACCESSION			
VERSION			
KEYWORDS			
EST			
BQ732661			
AGENCY			
IMAGE:5569808 5', mRNA sequence.			
BQ732661.1			
GI:21871558			
EST			
BQ732661			
878 bp			
mRNA			
linear			
EST 16-JUL-2002			
Xenopus laevis cDNA clone			



libraries" /lab\_host="DHL0B" /note="Vector: pSPOR1; Site\_1: SalI; Site\_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source cDNA libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

Alignment Scores:	4.09e-124	Length:	637
Pred. No.:		Matches:	208
Score:	119.00	Conservative:	1
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	99.52%	Indels:	0
Query Match:	11.20%	Gaps:	0
DR:	12		

US-09-964-956-13 (1-1896) x BG086250 (1-627)

Qy	1642	ThrProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHisGluHisGly	1661
Db	1	ACCCTGACCTGAAAGCGGGTCAAGCTTTGGCATCTGCTGAAGAACCAACATGAACATGGA	60
Qy	1662	AspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLeuThrArgLeu	1681
Db	61	GACCAGAAGGAGGGGACCGGGGAGCAAGATGGTGTCTGAGATCTATCTGCACGACATA	120
Qy	1682	LeuAlaThrLysGlyThrLeuGlnLysPheValAspAspLeuPheGluThrIlePheSer	1701
Db	121	TTGGCCACAAAGGCGACACTGCAGAAAGTTTGTGGATGACCTATTTCGAGACCATCTTCAGT	180
Qy	1702	ThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuAsp	1721
Db	181	ACGGCCACACCGTGGCTCGCTCTCCGCTAGCCATCAAGTACATGTTTGACTTCCTCGAT	240
Qy	1722	GluGlnAlaAspLysHisGlyIleHisAspProHisValArgHisThrTrpLysSerAsn	1741
Db	241	GAGCAGGCTGATAGCACCGCATTCATGACCCGACGTCGCCACACCTGGAGAGCAAC	300
Qy	1742	CysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheValPheAspIle	1761
Db	301	TGCCTGCCCTACGGTTTGGTGAATATGATCAAGAACCCCTCAGTTGTGTTTGACATC	360
Qy	1762	HisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnThrPheMetAspSer	1781
Db	361	CATAAGAACAGCATCACGGATGCCTGCTCTCTGTGGGGCCAGACCTTCATGACTCC	420
Qy	1782	CysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAla	1801
Db	421	TGCTCCACTTCGGACACCGACTGGGCAAGGATCCCCCTCCACAACGCTGCTCTATGCC	480
Qy	1802	LysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSerAspIleClyLysMet	1821
Db	481	AAGGACATCCCCAGCTACAAGAACCTGGGTAGAGAGGTATTATTTCAGACATTGGGAAGATG	540
Qy	1822	ProAlaIleSerAspClnAspMetAsnAlaTyrLeuAlaGluGlnSerArgMethHisMet	1841
Db	541	CSGGCAATAAGTCACAGACATGAATCATACCTGGCCGACGATCGCGGATGCGCATG	600

Qy	1842	AsnClupeaSnThrMetSerAlateu	1850
Db	601	AATGAGTTCAATACGATGAGTCGCTC	627
RESULT	9		
BI905192			
LOCUS	603167521P1	NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:525524 5',	linear EST 16-OCT-2001
DEFINITION		mRNA sequence.	
ACCESSION	BI905192		
VERSION	BI905192.1	GI:16167683	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		
	CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima		
	Bonaldo, Ph.D.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAM11645 row: j column: 05		
	High quality sequence start: 2		
	High quality sequence stop: 797.		

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

FEATURES	Location
source	1 971

1. .971  
/organism="Mus musculus"  
/strain="CZECH II"  
/db\_xref="taxon:10090"  
/clone\_image="5255524"  
/clone\_lib="NCI CGAP Lu33"  
/tissue\_type="pooled lung tumors"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Site\_1: NotI; Site\_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5]. TGTACCAATCTGAAGTCGGCGCCCTGTTTTTTTTTTT 3'1. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Ronaldo. "

BASE COUNT	244 a	262 c	254 q	211 t
100	100	100	100	100
200	200	200	200	200
300	300	300	300	300
400	400	400	400	400
500	500	500	500	500
600	600	600	600	600
700	700	700	700	700
800	800	800	800	800
900	900	900	900	900
1000	1000	1000	1000	1000

**BASE C  
ORIGIN**

Alignment Scores:	1.27e-123	Length:	971
Pred. No.:	1118.00	Matches:	214
Score:	87.55%	Conservative:	25
Percent Similarity:	78.39%	Mismatches:	31
Best Local Similarity:	11.19%	Indels:	3
Query Match:	13	Gaps:	0
Gap:	13		

USS-09-964-956-13 (1-1896) X BI905192 (1-971)

[illegible]

QY 1662 AspGlnLysGluGlyAspArgGlySerLysMetValSerGluLeuThrArgLeu 1681  
 Db 121 GATCACCAGAGAGAGACCGCGCAGCAAGATGGTCTCAGAAATATATCTCAAGCGCTG 180  
 QY 1682 LeuAlaThrLysGlyThrLeuGlnLysPheValAspLeuPheGluThrIlePheSer 1701  
 Db 181 CTGGCCACCAAGGGCAGCTTGCAGAAAGTTGTAGATGCTTTGAAACTGTGTTCAGT 240  
 QY 1702 ThrAlaHisArgGlySerAlaLeuProLeuAlaLeuLysTyrMetPheAspPheLeuAsp 1721  
 Db 241 ACAGCCACCGGGCTCAGCCTTACCTTGGCCATCAAGTACATGTTTCACTTCCTGGAT 300  
 QY 1722 GluGlnAlaAspLysHisGlyLysIleHisAspProHisValArgHisThrTrpLysSerAsn 1741  
 Db 301 GAACAGGCTGACAGCGCCAGATCAGTACCTGACCTGATGTCGTCACACCTGGAAGAGCAAC 360  
 QY 1742 CysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheValPheAspIle 1761  
 Db 361 TGTCTTACCTCTCGGTTCTGGGTGAATGTGATCAAGAATCCGCAATTTGTGTTGACATC 420  
 QY 1762 HisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnPheMetAspSer 1781  
 Db 421 CATAGATAGATCATCACAGACCGCTGTTTGTCTAGTGGTGGCCAGACCTTCATGGACTCC 480  
 QY 1782 CysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAla 1801  
 Db 481 TGTCTTACATCAGACGCGCTGGCAAGGACTCACCTTCCAACAAGCTCCTCTATGCC 540  
 QY 1802 LysAspIleProSerTyrLysAsnTrpValGluArgTyrTyrSerAspIleGlyLysMet 1821  
 Db 541 AAGGATATTCCCAATTACAAAGAGCTGGGTGGAGAGGTACTATCGAGATATCGCAAGATG 600  
 QY 1822 ProAlaIleSerAspGlnAspMetAsnAlaTyrLeuAlaGluGlnSerArgMetHisMet 1841  
 Db 601 GCATCCATCAGTACGACGAGCATGAGCGCTTACTTAGTGGAGAGTCCCGCTCCATGCT 660  
 QY 1842 AsnGluPheAsnThrMetSerAlaLeuSerGluIlePheSerTyrValGlyLysTyrSer 1861  
 Db 661 AATGACTTCAATGCTCTAAGTCACTCAGCGAGCTCTACTTCTATGTACCAAGTACCGT 720  
 QY 1862 GluGluIleLeuGlyProLeuAspHisAspAspGlnCys-GlyLysGlnLysLeuAlaTyr 1881  
 Db 721 CAGGAGATCCTCAGTCTCGTGGACCGAGATGCTCTTGTCCAGGAGCAACAAGCTTCGACA 780  
 QY 1881 r-LysLeuGluGlnValIle-ThrLeuMetSer 1891  
 Db 781 AGAAGCTGGAGCAGATCATTCACCTTGGTGTCC 813

RESULT 10  
 BQ723197  
 LOCUS  
 DEFINITION AGENCOURT\_8470039 Lupski\_sympathetic\_trunk Homo sapiens cDNA clone  
 IMAGE:6195860 5', mRNA sequence.  
 ACCESSION BQ723197  
 VERSION  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 918)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: c9abps-remail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
 Plate: LLAM13603 row: f column: 21  
 High quality sequence stop: 692.  
 FEATURES  
 source 1..918  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Lupski\_sympathetic\_trunk"  
 /sex="male"  
 /tissue\_type="sympathetic trunk"  
 /dev\_stage="adult, 16 yr"  
 /lab\_host="DH10B"  
 /note="vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
 NOTI; Site\_2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCACCGGTCG-3' and  
 5'-GACTAGTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.9 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine); available through Life  
 Technologies."  
 BASE COUNT 225 a 266 c 259 g 167 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.86e-122 Length: 918  
 Score: 1108.00 Matches: 231  
 Percent Similarity: 77.78% Conservative: 28  
 Best Local Similarity: 69.37% Mismatches: 33  
 Query Match: 11.09% Indels: 41  
 DB: 14 Gaps: 3  
 US-09-964-956-13 (1-1896) x BQ723197 (1-918)  
 QY 1415 LysAsnLeuGluSerLysAsnHisProLysLeuLeuArgThrGluSerValAla 1434  
 Db 3 AGAAACCTGGAGAACAGAACCCACCCAAAGCTGTACTCCGAGGACAGAGTCTGGCT 62  
 QY 1435 GluLysMetLeuThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAla 1454  
 Db 63 GAAAGATGCTGACCAATTGGTTCGCTTCTCTGCAAGATTCCTTAAAGAGGTGCGCA 122  
 QY 1455 GlyGluProLeuPheSerLeuPheCysAlaIleLysGlnMetGluLysGlyProIle 1474  
 Db 123 GGGAGCCACTCTTCATCTATATCTGTCCATCAAGCAGCAGATGGAGAGGCCCCATT 182  
 QY 1475 AspAlaIleThrGlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGln 1494  
 Db 183 GATGCCATCAGCGGAGGCGCGCTACTCTCCAGGAGGACAAGCTCATCCGCGCAGCAG 242  
 QY 1495 IleAspTyrLysThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGlu 1514  
 Db 243 ATCGAGTACAGAACCCCTGATCTCACTGCTCAACCTGACACGAGGAGGAGGAGTCCAG 302  
 QY 1515 ValProValLysIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAsp 1534  
 Db 303 ATCCCACTGAAGGTGTTAAACTGTGACACCATCACACAGGTCAAGGAGAGATCTTCAT 362  
 QY 1535 AlaIlePheLysAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeuGlu 1554  
 Db 363 GCGGTGTATAAGAAATGTGCCCTATTCCAGCGGCGGAGGAGTGGACATGGACATTGGAG 422  
 QY 1555 TrpArgGlnGlySerGlyAlaArgMetIleLeuGlnAspGluAspIleThrThrLysIle 1574  
 Db 423 TGGCGCCAAAGCCGGATCGCCCGGTGCTGTGCAAGATGAGGACATCACCACCAAGATT 482  
 QY 1575 GluAsnAspTrpLysArgLeuAsnThrLeuAlaHisTyrGlnValProAspGlySerVal 1594  
 Db 483 GAGGCTGACTGGAAGCGGCTCAACACACTGATGCTATATCATGCTGTACAGAGGTGGTG 542  
 QY 1595 ValAlaLeuValSerLysGlnValThrAlaTyrAsnAlaValAsnSerThrValSer 1614

```

Db      543 GTGGCTCTGGTCCCAACAGAGACTCTCTCTACAACATCCCTCCCTCTGCGAGATCTCC 602
QY      1615 ArgThrSerAlaSerLysTyrGluAsnMetIleArgTyrThrGlySerProAspSerLeu 1634
Db      603 CGGACGTCCATCAGCAGATACGACTCTCTCTCAGGTATACGGGCGACCCCGACAGCTG 662
QY      1635 ArgSerArgThrProMetIleThrProAspLeuGluSerGlyValLys-MetTrpHisLe 1654
Db      663 CGGTCCGGGGCCCGATGATCACCAGAGACTGGAAGTGGGTCAAGGGTGTGGCATCT 722
QY      1654 u-VallYsnHisGluHis-GlyAspGlnLysGluGlyAspArgGlySer-LysMetVa 1673
Db      723 GGGTGAAGAACCATGACGACCGGTGACCAAGAGAGGTGACCGGGGAGCAAGATGGT 782
QY      1673 l-SerGluIleTyr-LeuThrArgLeuAla---ThrLysGlyThr-LeuGlnLys-- 1690
Db      783 GTTCCGAGATCTACCTGACCGGGCTACTGGGCCACCCAGGCGACCCCTGTCAAAAAG 842
QY      1691 -PheValAspLeuPheGluThrIlePheSerThrAlaHisArgGlySerAlaLeuPr 1710
Db      843 TTTTGTGGACGACTT----- 858
QY      1710 oLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAlaAspLysHisGlyIleHi 1730
Db      859 -----GGTTTGA 866

QY      1730 sAspProHisValArgHisThrTrp 1738
Db      867 AGACCCCTGGTTCAGGACACCTGTGG 891

RESULT 11
BQ434703 844 bp mRNA linear EST 24-MAY-2002
LOCUS AGENCOURT_7909800 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6156354
DEFINITION 5', mRNA sequence.
ACCESSION BQ434703
VERSION BQ434703.1 GI:21173779
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 844)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue procurement: ATCC
CDNA Library preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13500 row: h column: 19
High quality sequence stop: 597.
FEATURES
Location/Qualifiers
1..844
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6156354"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT 202 a 256 c 246 g 140 t
ORIGIN
Alignment Scores:

```

```

Pred. No.: 4.5e-121 Length: 844
Score: 1096.00 Matches: 205
Percent Similarity: 91.94% Conservative: 23
Best Local Similarity: 82.66% Mismatches: 19
Query Match: 10.97% Indels: 1
DB: 14 Gaps: 0

US-09-964-956-13 (1-1896) x BQ434703 (1-844)
QY 1648 GlyValLysMetTrpHisLeu-VallYsnHisGluHisGlyAspGlnLysGluGlyAs 1667
Db 3 GGCACCAAGCTGTGGACCTGGTGAAGAACCAACCAACCACTGGACCAAGCGGTGAGGTGA 62
QY 1667 pArgGlySerLysMetValSerGluIleTyrLeuThrArgLeuLeuAlaThrLysGlyTh 1687
Db 63 CCGCGCGACAGATGGTCTCGAGACTTCTTGACACGGCTACTGTGCCACCAAGGCAC 122
QY 1687 rLeuGlnLysPheValAspLeuPheGluThrIlePheSerThrAlaHisArgGlyLys 1707
Db 123 ACTGCAGAAAGTTTGTGGACGACCTGTTGAGACCACTTTTCAGCACCGCACACCGGGCTC 182
QY 1707 rAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuAspGluGlnAlaAspLysHi 1727
Db 183 AGCCCTGGCGCTGGCCATCAAGTACATGTTGCACTTCTTGGATGAGCAGCGCCAGCAAGCA 242
QY 1727 sGlyIleHisAspProHisValArgHisThrTrpLysSerAsnLysLeuProLeuArgPh 1747
Db 243 CCAGATCCACGATGCTGTGAGCGCCACACCTCGAAGAGCAACTGCTGCCCTGCCCTT 302
QY 1747 eTrpValAsnMetIleLysAsnProGlnPheValPheAspIleHisLysAsnSerIleTh 1767
Db 303 CTGGGTGAACGATGATCAAGAACCCACAGTTGTGTCGACATTCACAAGAACAGCATCAC 362
QY 1767 rAspAlaCysLeuSerValValAlaGlnThrPheMetAspSerCysSerThrSerGluHi 1787
Db 363 GGACGCGCTGTTGTCGGTGGTGGCCACACCTCGAAGAGCAACTGCTGCCCTGCCCTT 422
QY 1787 sArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAlaLysAspIleProSerTy 1807
Db 423 CAAGCTGGGCAAGGACTCACCTCCCAACAGCTGCTCTAGCCCAAGCAGCATCCCCAACA 482
QY 1807 rLysAsnTrpValGluArgTyrTyrSerAspIleGlyLysMetProAlaIleSerAspGl 1827
Db 483 CAAGAGCTGGTGGAGAGGTACTATGACAGACATCGCCCAAGATGACGACCATCAGCGACA 542
QY 1827 nAspMetAsnAlaTyrLeuAlaGluGlnSerArgMetHisMetAsnGluPheAsnThrMe 1847
Db 543 GGACATGAGTGCCTATCTGCTGAGCAGCTCCCGCTGACGACCATGAGATGAGATGAGATG 602
QY 1847 tSerAlaLeuSerGluIlePheSerTyrValGlyLysTyrSerGluGluLeuLeuGlyPr 1867
Db 603 GAGCGCTTGCACGAGATCTACTCTACATCCCAAGATACAAAGATGATGATGATGATGATG 662
QY 1867 oLeuAspHisAspAspGlnCysGlyLysGlnLysLeuAlaTyrLysLeuGluGlnValIl 1887
Db 663 CCTGGAGAAAGATGAGCAGCGCGCGCGAGCGCTGCGGAGCAAGCTGGAGCAGGTGTGT 722
QY 1887 eThrLeuMetSerLeuAspSer 1894
Db 723 GGACACGATGGCCCTGAGCAGC 744

RESULT 12
BQ878090 851 bp mRNA linear EST 16-AUG-2002
LOCUS AGENCOURT_8072266 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6089613
DEFINITION 5', mRNA sequence.
ACCESSION BQ878090
VERSION BQ878090.1 GI:22270098
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```



```

/lab_host="xl-2blue"
/note="vector: pAMP1; Site_1: NotI; Site_2: SalI"
BASE COUNT 173 a 225 c 228 g 136 t
ORIGIN
Alignment Scores:
Pred. No.: 7,45e-119 Length: 762
Score: 1077.00 Matches: 213
Percent Similarity: 90.31% Conservative: 20
Best Local Similarity: 82.56% Mismatches: 21
Query Match: 10.78% Indels: 5
DB: 9 Gaps: 1

US-09-964-956-13 (1-1896) x ALL34739 (1-762)
QY 1296 LeuGlnThrAspIleHisGluLeuThrSerAspLeuAspGlyAlaGlyIleProPheLeu 1315
Db 1 CTGCAGACAGACATCCACGAGCTGACCAATGACCTGGAGCGTGGCGGATCCCTTCCTT 60
QY 1316 AspTyrArgThrTyrThrMetArgValLeuPheProGlyIleGluAspHisProValLeu 1335
Db 61 GACTACCGGACATATGCCATGCGGTGCTCTTCTGGGATCGAGGACCCCTGTGCTC 120
QY 1336 ArgAspLeuGluValProGlyTyrArgGlnGluArgValGluGlyLeuLysLeuPhe 1355
Db 121 AAGGAGATGGAGGTG-----CAGGCCAATGTGGAGAAGTGGCTGACACTGTTC 168
QY 1356 AlaGlnLeuIleAsnAsnLysValPheLeuLeuSerPheIleArgThrLeuGluSerGln 1375
Db 169 GGCAGCTGCTGACCAAGAAGCAGCTTCCTGCTGACCTTCATCGCAGCTGGAGGCAG 228
QY 1376 ArgSerPheSerMetArgAspArgGlyAsnValAlaSerLeuIleMetThrValLeuGln 1395
Db 229 CGCAGCTTCTCCATGCGGACCGCGGAATGGCCCTGCTCATCATGAGCGCCCTGCAG 288
QY 1396 SerLysLeuGluTyrAlaThrAspValLeuLysGlnLeuLeuAlaAspLeuIleAspLys 1415
Db 289 GCGGAGATGGAATACGCCACAGCGTGTCTCAAGCAGCTGCTTCCGACCTTCATCGAGA 348
QY 1416 AsnLeuGluSerLysAsnHisProLysLeuLeuLeuArgArgThrGluSerValAlaGlu 1435
Db 349 AACCTGGAGAGCAAGAACCAACCCAGCTGCTACTGCGCGGAGTGAAGAGGCCCATTTGAC 408
QY 1436 LysMetLeuThrAsnTrpPheThrPheLeuLeuTyrLysPheLeuLysGluCysAlaGly 1455
Db 409 AAGATGCTAACTAAGTGTTCACCTTCTCTTATTAAGTTCCTCAAGGAGTGGCTGGG 468
QY 1456 GluProLeuPheSerLeuPheCysAlaIleLysGlnMetGluLysGlyProIleAsp 1475
Db 469 GAGCGCGTGTTCATGCTGTACTGCGCATCAAGCAGCAGATGGAGAAGGCCCATTTGAC 528
QY 1476 AlaIleThrGlyGluAlaArgTyrSerLeuSerGluAspLysLeuIleArgGlnGlnIle 1495
Db 529 GCATACACGGGTGAGGACCGCTACTCCTCGAGTGAAGCAAGCTCATCCGCGCAGCAT 588
QY 1496 AspTyrLysThrLeuValLeuSerCysValSerProAspAsnAlaAsnSerProGluVal 1515
Db 589 GACTACAAGACATGACCCCTGAACCTGTGAACCTGAGAAATGAGAATGACCTGAGGTG 648
QY 1516 ProValLysIleLeuAsnCysAspThrIleThrGlnValLysGluLysIleLeuAspAla 1535
Db 649 CCGGTGAAGGGCTGGAGTGTGACACGGTTCACCCGAGCAAGGAGGAGAGCTGCTGGACGCT 708
QY 1536 IlePheLysAsnValProCysSerHisArgProLysAlaAlaAspMetAspLeu 1553
Db 709 GCCTACAAGGCGGTGCCTAC-TCCACGCGGCCCAAGGCCGCGGACATGGACCTG 761

RESULT 14
BG294786 742 bp mRNA linear EST 21-FEB-2001
LOCUS 602391607F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4503720 5',
DEFINITION mRNA sequence.
ACCESSION BG294786
```

```

BG294786.1 GI:13055976
EST.
house mouse.
Mus musculus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 742)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc. (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0374 row: m column: 01
High quality sequence stop: 712.
Location/Qualifiers.
1..742
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4503720"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 187 a 207 c 182 g 166 t
ORIGIN
Alignment Scores:
Pred. No.: 2,86e-118 Length: 742
Score: 1072.00 Matches: 211
Percent Similarity: 91.53% Conservative: 16
Best Local Similarity: 85.08% Mismatches: 21
Query Match: 10.73% Indels: 3
DB: 12 Gaps: 0

US-09-964-956-13 (1-1896) x BG294786 (1-742)
QY 1562 ArgMetIleLeuGlnAspGluAspIleThrLysIleGluAsnAspTrpLysArgLeu 1581
Db 1 AGAGTGTGTGTGCGAGGAGAGACATTACCAAAATAGAGGGTGCCTGGAGCGGCTT 60
QY 1582 AsnThrLeuAlaHisTyrGlnValProAspGlySerValAlaLeuValSerLysGln 1601
Db 61 AACACACTGATGCATATCCAGGTGTGACACACATCCGTGGTGTCTGCTTCTCTAAGCAG 120
QY 1602 ValThrAlaTyrAsnAlaValAsnAsnSerThrValSerArgThrSerAlaSerLysTyr 1621
Db 121 ACCCTCCTCTACACATCCCTGCTGCGCAGCATCTCTCGGACATATTCATGACATAT 180
QY 1622 GluAsnMetIleArgTyrThrGlySerProAspSerLeuArgSerArgThrProMetIle 1641
Db 181 GACTCTTCTTCCAGGTACAGGACGCCAGCAGCAGCTCCCGTCCCGGGTCCCATGATC 240
QY 1642 ThrProAspLeuGluSerGlyValLysMetTrpHisLeuValLysAsnHisGluHisGly 1661
Db 241 ACCCCAGACTTGGAGCGGTGTCAAGTTTGGCATCTGTGAGAAATCATGACCATGGT 300
QY 1662 AspGlnLysGluGlyAspArgGlySerLysMetValSerGluIleTyrLeuThrArgLeu 1681
Db 301 GACCAGAGAGGAGGTGACCGGGGACCAAAATGGTGTCTGAGATCTACTTGACCGGCTT 360
QY 1682 LeuAlaThrLysGlyThrLeuGlnLysPheValAspLeuPheGluThrIlePheSer 1701
```

Db 361 CTAGCCACCAAGGGCCCTGCGAGAAATTTGTGGAGACATTTGTTGAGACCTTTGTTGAGC 420  
 QY 1702 ThrAlaHisArgGlySerAlaLeuProLeuAlaIleLysTyrMetPheAspPheLeuAsp 1721  
 Db 421 ACTGTGCACCGGGGTAGTCTCTCCCTAGCCATCAAGTACATGTTGATTTCTCTGGAT 480  
 QY 1722 GluGlnAlaAspLysHisGlyIleLeHisAspProHisValArgHisThrLysSerAsn 1741  
 Db 481 GAGCAGGCACACAGACAGTATCCACGACACAGATGTCGGGCACACCTGGAAAGCAAC 540  
 QY 1742 CysLeuProLeuArgPheTrpValAsnMetIleLysAsnProGlnPheValPheAspIle 1761  
 Db 541 TGGCTTCCACATCTGTTCTGGGGTGAATGTCATCAAGAACCTCAATTTGTTATTTGACATC 600  
 QY 1762 HisLysAsnSerIleThrAspAlaCysLeuSerValValAlaGlnThrPheMetAspSer 1781  
 Db 601 CAACAGGGCAGCATCAGACATGCTGCTCTCTGTGTAGCCAGACCTTTATGGACTCC 660  
 QY 1782 CysSerThrSerGluHisArgLeuGlyLysAspSerProSerAsnLysLeuLeuTyrAla 1801  
 Db 661 TGTTCACATCAGAGCAGCA-CTAGGCAAGGACTCACCCTCCCAACAG-CTGCTCTATGCC 718  
 QY 1802 LysAspIleProSerTyrLysAsn 1809  
 Db 719 AGGATAT-CCCAAGTATATAAGAAC 741

## RESULT 15

## LOCUS

BB656076 612 bp mRNA linear EST 26-OCT-2001  
 BB656076 RIKEN full-length enriched, 12 days embryo spinal ganglion  
 Mus musculus cDNA clone D130060F15 5', mRNA sequence.

## ACCESSION

VERSION BB656076.1 GI:16489904

## KEYWORDS

SOURCE house mouse.

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 612)

## REFERENCE

Arahawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,  
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda  
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Sasaki  
 Okazaki,Y., Okido,T., Saito,K., Sakai,C., Sakai,K., Sano,H., Sasaki  
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,  
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T.,  
 Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Mouse ESTs (Arahawa,T., et al. 2001)  
 Unpublished (2001)

## TITLE

## JOURNAL

## COMMENT

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh  
 M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi,K., Fujiwara,K., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura  
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
 Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara  
 Y. and Hayashizaki,Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamazaki,I., Aizawa  
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and  
 Hayashizaki,Y.  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.

## e mouse tissues.

## FEATURES

## source

Location/Qualifiers

1..612

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="D130060F15"

/clone\_lib="RIKEN full-length enriched, 12 days embryo

spinal ganglion"

/tissue\_type="spinal ganglion"

/dev\_stage="12 days embryo"

/lab\_host="DH10B"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGAGCGCGCACTCGAGTTTTTTTTTTTTTTTNN 3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. Second strand cDNA was prepared with the

primer adapter of sequence [5'

GAGAGAGAGATTCGAGTTAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA

was cleaved with BamHI and XhoI. Vector: a modified

pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 147 a 176 c 160 g 129 t

## ORIGIN

## Alignment Scores:

Pred. No.: 1,05e-117 Length: 612

Score: 1066.00 Matches: 203

Percent Similarity: 99.51% Conservatives: 0

Best Local Similarity: 99.51% Mismatches: 1

Query Match: 10.67% Indels: 0

DB: 10 Gaps: 0

US-09-964-956-13 (1-1896) x BB656076 (1-612)

QY 1043 ValArgIleGluProGluTrpSerIleValSerGlyAsnThrProIleAlaValTrpGly 1062  
 Db 1 GTGGGGATTGAACCAATGGAGCATTTGTCAGTGGGAACACACCATTTGCTGCTGGGA 60  
 QY 1063 ThrHisLeuAspLeuIleGlnAsnProGlnIleArgAlaLysHisGlyGlyLysGluHis 1082  
 Db 61 ACTCACCTGGACCTCATACAGAACCCAGATCCCTGCCAACATGGAGAAAGAACAC 120  
 QY 1083 IleAsnIleCysGluValLeuAsnAlaThrGluMetThrCysGlnAlaProAlaLeuAla 1102  
 Db 121 ATCAACATCTCTGAGGTCTTAAATGCTACAGAGATGACCTGCCAGGCTCCAGCCCTTGC 180  
 QY 1103 LeuGlyProAspHisGlnSerAspLeuThrGluArgProGluGluPheGlyPheLeu 1122  
 Db 181 CTGGGTCCGACCAACACCATGATCTCATGAGAGCCCTGGAAGAATTTGTTTCATCTCG 240  
 QY 1123 AspAsnValGlnSerLeuLeuIleLeuAsnLysThrAsnPheThrTyrTyrProAsnPro 1142  
 Db 241 GACATGTCCAGTCTCTCTCTAATTTCTCAACAGACCACTTCCACTATTTATCTACCCC 300  
 QY 1143 ValPheGluAlaPheGlyProSerGlyIleLeuGluLeuLysProGlyThrProIleIle 1162  
 Db 301 GTGTTTGGAGCTTTCAGTCCCTCGGAATCCTGGAGCTCAAGCCGCGCACCCCATTC 360  
 QY 1163 LeuLysGlyLysAsnLeuIleProValAlaGlyGlyAsnValLysLeuAsnThrThr 1182

Mon Jun 30 08:38:11 2003

Db 361 CTRAGGSCAAGACTTGATCCGCGCTGTGGCCGAGGCAATGTGAAGCTGAACACACG 420  
Qy 1183 ValLeuValGlyGluLysProCysThrValThrValSerAspValGlnLeuLeuCysGlu 1202  
Db 421 GTGCTGCTTGGGGAGAAACCATGCACCGTGACAGTACAGATGTGCAGCTGCTCTGCCAG 480  
Qy 1203 SerProAsnLeuIleGlyArgHisLysValMetArgValGlyGlyMetGluTyrSer 1222  
Db 481 TCTCCCAACCTCATCGGCAGCAGAAAGTGATGGCAGGGTGGGTGGCATGCAGTATGCC 540  
Qy 1223 ProGlyMetValTyrIleAlaProAspSerProLeuSerLeuProAlaIleValSerIle 1242  
Db 541 CCTGGGATGGTATACATGCCCCAGACAGCCCGCTCAGCCGCGCCGCAATGTTACGATC 600  
Qy 1243 AlaValAlaGly 1246  
Db 601 GCGGTGGCTGGG 612

Search completed: June 28, 2003, 15:57:58  
Job time : 5438 secs